3-1-2019

Genetics of group A streptococci

Kyu Hong Cho
Indiana State University

Gary C Port
Washington University School of Medicine in St. Louis

Michael Caparon
Washington University School of Medicine in St. Louis

Microbiology Spectrum
American Society for Microbiology Press

Genetics of Group A Streptococci

KYU HONG CHO,[1] GARY C. PORT,[2*] and MICHAEL CAPARON[2]

[1]Department of Biology, Indiana State University, Terre Haute, IN 47809; [2]Department of Molecular Microbiology, Washington University School of Medicine, Saint Louis, MO 63110

ABSTRACT *Streptococcus pyogenes* (group A streptococcus) is remarkable in terms of the large number of diseases it can cause in humans and for the large number of streptococcal factors that have been identified as potential virulence determinants for these diseases. A challenge is to link the function of potential virulence factors to the pathogenesis of specific diseases. An exciting advance has been the development of sophisticated genetic systems for the construction of loss-of-function, conditional, hypomorphic, and gain-of-function mutations in targeted *S. pyogenes* genes that can be used to test specific hypotheses regarding these genes in pathogenesis. This will facilitate a mechanistic understanding of how a specific gene function contributes to the pathogenesis of each streptococcal disease. Since the first *S. pyogenes* genome was completed in 2001, hundreds of complete and draft genome sequences have been deposited. We now know that the average *S. pyogenes* genome is approximately 1.85 Mb and encodes ~1,800 genes and that the function of most of those genes in pathogenesis remains to be elucidated. However, advances in the development of a variety of genetic tools for manipulation of the *S. pyogenes* genome now provide a platform for the interrogation of gene/phenotype relationships for individual *S. pyogenes* diseases, which may lead to the development of more sophisticated and targeted therapeutic interventions. This article presents an overview of these genetic tools, including the methods of genetic modification and their applications.

INTRODUCTION

Streptococcus pyogenes (the group A streptococcus) is remarkable in terms of the large number of very different diseases it can cause in humans. These range from superficial and self-limiting diseases of the pharynx (e.g., pharyngitis, commonly known as strep throat) and skin (impetigo) to infections that involve increasingly deeper layers of tissue and are associated with increasing degrees of destruction of tissue (e.g., erysipelas, cellulitis, necrotizing fasciitis, and myositis). The organism has the ability to spread rapidly through tissue and to penetrate into the vasculature to cause lethal sepsis. Other diseases result from the production of toxins that spread through tissue or spread systemically from a site of local infection (scarlet fever and toxic shock syndrome). Still other diseases are the result of an immunopathological response on the part of the host that is triggered by a streptococcal infection. These diseases include rheumatic fever, acute glomerulonephritis, certain types of psoriasis, and potentially even some forms of obsessive-compulsive syndrome disorder (pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections).

S. pyogenes is even more remarkable in terms of the large number of factors that have been identified as potential virulence determinants for these various diseases. These include surface proteins (M proteins, fibronectin-binding proteins, surface dehydrogenases, C5a peptidase), the hyaluronic acid capsule, secreted degradative enzymes (several distinct DNases, a cysteine protease, NAD^+-glycohydrolase, hyaluronidases,) and many secreted toxins (streptolysin S, streptolysin O, the pyro-

Received: 4 December 2018, **Accepted:** 8 January 2019, **Published:** 1 March 2019

Editors: Vincent A. Fischetti, The Rockefeller University, New York, NY; Richard P. Novick, Skirball Institute for Molecular Medicine, NYU Medical Center, New York, NY; Joseph J. Ferretti, Department of Microbiology & Immunology, University of Oklahoma Health Science Center, Oklahoma City, OK; Daniel A. Portnoy, Department of Molecular and Cellular Microbiology, University of California, Berkeley, Berkeley, CA; Miriam Braunstein, Department of Microbiology and Immunology, University of North Carolina-Chapel Hill, Chapel Hill, NC, and Julian I. Rood, Infection and Immunity Program, Monash Biomedicine Discovery Institute, Monash University, Melbourne, Australia

Citation: Cho KH, Port G, Caparon M. 2019. Genetics of Group A *Streptococci*. *Microbiol Spectrum* 7(2):GPP3-0056-2018. doi:10.1128/microbiolspec.GPP3-0056-2018.

Correspondence: Michael Caparon, caparon@wustl.edu

*Current address: Elanco Animal Health, Natural Products Fermentation, Eli Lilly and Company, Indianapolis, IN 46285

genic exotoxins, streptococcal superantigens, streptokinase). This is only a partial list of the potential virulence factors this bacterium can produce, and many of these are considered in more detail elsewhere.

Until recently, the function that any of these potential virulence factors contributed to the pathogenesis of any specific streptococcal disease was only poorly understood. A major reason for this deficiency was the lack of sophisticated genetic systems that could be applied to the analysis of a specific virulence factor according to modern molecular criteria. These criteria were succinctly outlined by Falkow in his "Molecular Koch's Postulates" (1). These state that (i) the phenotype under investigation should be associated with pathogenic members of a species, (ii) the gene(s) associated with the virulence trait should be identified and isolated by molecular methods, (iii) specific inactivation of identified gene(s) should lead to a measurable loss in pathogenicity, and (iv) reintroduction of the unmodified wild-type gene should lead to a restoration of pathogenicity. An exciting development is that, over the past several decades, the work of many groups has contributed to the development of sophisticated group A streptococcal genetic systems that now allow the many diseases that this organism can cause to be studied at this level of molecular resolution. The modern *S. pyogenes* genetic modification systems are supported by the numerous complete genome sequences now available. Since the first complete genome sequence of an *S. pyogenes* strain in 2001 (2), 61 complete genome sequences and approximately 300 draft sequences have been deposited as of March 2018 (3). The average *S. pyogenes* genome is approximately 1.85 Mb and encodes ~1,800 genes, the function of many of which remains to be elucidated (4). The other major advance has been in the development of a variety of genetic tools aimed at manipulating the *S. pyogenes* genome. The goal of this article is to present an overview of these tools, including the methods of genetic modification and their applications.

GENETIC EXCHANGE

Transduction

A key element of any genetic system involves some system of genetic exchange between bacterial hosts, allowing the construction of an altered genome in the target host which can then be subjected to an analysis of its virulence phenotypes. Unlike several other species of streptococci, the group A streptococci are not naturally competent for the uptake of exogenous DNA, even though their genomes encode many of the genes required for transformation (see below). Conjugative DNA transfer does occur in group A streptococci; however, this is restricted to the transfer of conjugative plasmids and conjugative transposons (see below), and there is no evidence of mobilization of chromosomal markers. There is also no evidence that any important virulence traits are encoded by these types of mobile elements, although they are undoubtedly important in the transmission of resistance to various antibiotics.

Even though there is no evidence to support genetic exchange by transformation or by conjugation, analysis of several polymorphic loci, most notably the genes which encode the M proteins and the fibronectin- and collagen-binding proteins and T antigen (FCT)/pilus region have provided considerable evidence of horizontal transfer of genetic material among natural populations of *S. pyogenes* (5). In this regard, considerable attention has been focused on the contribution of phage to genetic transfer (5). *S. pyogenes* strains are rich in phage, and several of these have been demonstrated to encode virulence factors, including the pyrogenic exotoxins SPE-A and SPE-C (6). Thus, it is not surprising that transduction was the first mechanism of genetic exchange that was exploited in the manipulation of the *S. pyogenes* genome (7). The most highly developed of these are derivatives of phage A25. This lytic phage is classified as a Bradley group B phage that recognizes the peptidoglycan of groups A, C, and G streptococci as its cellular receptor. It has a 35-kb double-stranded genome with circular permutation and terminal repetition (8), and it is proficient in transducing markers *in vitro* (for a detailed description of a method of transduction, see reference 9). The most useful derivatives of A25 for use in transduction are those developed by Malke and colleagues, who constructed a derivative with two distinct temperature-sensitive lesions ($A25_{ts1-2}$) that becomes defective for growth at 37°C (10). This feature is useful for transduction, since it allows the production of a transducing lysate of the donor host at the permissive temperature (30°C) but prevents the killing of transduced hosts when infection is performed at the nonpermissive temperature. The A25 phage was used in some of the first mutagenesis studies in group A streptococci, including the original identification of Mga, the master regulator of M protein (7). In spite of the early use of this method, genetic manipulation via phage transduction is inherently limited in application since it requires preexisting antibiotic resistance markers within or near genes of interest. Additionally, some phage possess narrow host specificities, thereby limiting their utility to only a few strains. Furthermore, phage transduction can be laborious. Fortunately, simpler and more rapid

genetic engineering methods, including electroporation of shuttle plasmids and saturating transposon mutagenesis, have largely replaced phage transduction (see below).

Transformation

One of the most significant advances in genetic technology for *S. pyogenes* has been the development of methods for transformation. This has allowed the introduction of heterologous DNA into an *S. pyogenes* host and has opened the door to a large number of techniques for mutagenesis and allelic exchange that are described in greater detail below. The breakthrough technology that began the era of transformation for *S. pyogenes* was the introduction by several companies of reasonably priced instruments for the introduction of DNA by electroporation. As is true for most methods of transformation, success with electroporation-based transformation results from careful attention to growth conditions. Electroporation of streptococci generally requires that the cell walls be weakened. Most successful methods have adapted the techniques originally developed by Dunny and colleagues for electroporation of the enterococci (11). This method uses cells from the early exponential stages of growth in medium supplemented with glycine. The addition of glycine is thought to contribute to a decreased level of cross-linking in the cell wall, and the exact stage of growth and concentration of glycine is determined empirically (for a detailed description of the method, see references 9, 12–14).

It is of interest to note that the genomic sequences of several *S. pyogenes* strains of the M1, M3, and M5 serotypes have revealed that while they have genes required for DNA uptake by natural transformation and that while these genes are expressed in response to a peptide pheromone, they remain nontransformable (15). In addition, a cluster of genes associated with the ability to cause disseminated infection, called the Sil locus, contains genes highly homologous to the *comAB* competence locus (16). To date, the Sil genes have only been found among serotype M14 and M18 strains, and since their G+C content is markedly lower than the rest of the *S. pyogenes* chromosome, they may have been more recently acquired by these lineages from horizontal transmission (17).

Plasmid Technology

The development of efficient transformation systems has allowed for the rapid genetic modification of *S. pyogenes* hosts through the introduction of plasmids into *S. pyogenes*. Applications include (i) construction of chromosomal allelic replacements, gene deletions, and site-specific mutations; (ii) complementation of chromosomal mutations *in trans*; (iii) analysis of reporter gene expression; and (iv) delivery of transposable elements to construct random mutant libraries for forward genetic screens. Common Gram-negative plasmid replicons (including pUC, pBR322, and pSC101) are unable to replicate in Gram-positive species and can therefore serve as "suicide" delivery vectors for transposon delivery or recombination-based chromosomal integrative elements (see below). Also, many of the plasmids which have proven useful for genetic analysis in other Gram-positive species, including pC194, pE194, and pUB110, appear unable to replicate reliably in *S. pyogenes*. Fortunately, a handful of plasmids have been identified that do replicate in *S. pyogenes*, and these plasmids have been engineered over the past 20 years to give rise to a growing set of tools aimed at manipulating the *S. pyogenes* chromosome (Table 1).

The most widely used plasmid for *S. pyogenes* genome engineering is derived from pWV01 (18). This plasmid was isolated from a lactococcal species and is the prototype member of a family of "promiscuous replicons" that have the remarkable property that they can replicate in both Gram-negative (e.g., *Escherichia coli*) and Gram-positive (e.g., *S. pyogenes*) hosts. pWV01 replicates using a rolling-circle type mechanism characteristic of many single-stranded bacteriophages and likely exists in a single-stranded conformation for extended periods of time in replicating cells. This can have consequences for plasmid structural integrity, segregation instability, and expression of open reading frames inserted for complementation or ectopic or orthogonal expression (19). Common derivatives of pWV01 include plasmids such as pABG5 (Fig. 1), which include an *S. pyogenes* promoter for ectopic or orthogonal expression (20), and the high-copy variant pBAV1kT5, which lacks a set of inverted repeats and an open reading frame thought to be responsible for copy number regulation (21). A versatile and widely used derivative of pWV01 known as pG+host4 was developed by Maguin and coworkers to contain a temperature-sensitive (ts) replicon (22). The plasmid contains four amino acid substitutions in the RepA protein, which is responsible for nicking one DNA strand at the plus origin to initiate replication, thus rendering the process temperature sensitive. Subsequent derivatives of this plasmid include pJRS233 (23) and pGCP213 (24), which add low copy (pSC101) or high copy (ColE1/pUC) thermo-stable Gram-negative replication origins, respectively, to facilitate various applications for subcloning and plasmid propagation in *E. coli*. In addition, pGCP213 includes a full-length gene encoding LacZ containing the multi-

TABLE 1 Common plasmids utilized for *S. pyogenes* genome engineering

Capability	Name	Ori G+/G−	Ab cassette(s)	Features	Reference(s)
Site-directed mutagenesis	pJRS233	pWV01 ts/pSC101	*ermC*	Low copy (*E. coli*) thermosensitive shuttle vector	23
	pGCP213	pWV01 ts/pUC (pBR322)	*ermC*	High copy (*E. coli*) thermosensitive shuttle vector	24
	pCRK	pWV01 ts/pUC (pBR322)	*aphA3*	High copy (*E. coli*) thermosensitive shuttle vector	25
	pCRS	pWV01 ts/pUC (pBR322)	*aad9*	High copy (*E. coli*) thermosensitive shuttle vector	25
	pHY304	pWV01 ts/NA	*ermC*	Thermosensitive vector, must be grown at 30°C	111
	pSinS/pHlpK	pCRS (ΔrepA) pCRK (ΔpUC)	*aad9*/*aphA3*	Conditional suicide plasmid pSin stabilized by temperature-sensitive helper plasmid	42
Complementation	pABG5	pWV01/NA	*aphA3*/*catA*	Thermostable shuttle vector, ProfA promoter	20
	p7INT, p7INT[a] (pGCP365)	NA/pUC (pBR322)	*ermC*	Integrates into prophage tmRNA attB site	33, Port, Caparon Unpublished
	pEU7742	pAMβ1/NA	*aad9*	Single copy with toxin/antitoxin, contains theophylline-inducible riboswitch and *gusA* reporter	32
	pGCP189	pWV01 ts/pSC101	*ermC*	Insert gene of interest downstream of GuaB in synthetic polycistron	45
Transposon	pMGC57-TnErm	NA/pHSG396 (pBR322)	*catA*/*ermC* (for Tn)	Tn*4001*	80
	EZ-Tn*5*	NA/NA	*ermC*	Tn*5*	43
	pOSKAR	pWV01 ts/pUC (pBR322)	*aad9*/*aphA3* (for Tn)	Mariner Kan-Tn flanked by T7 For microarrary TraSH	25
	pKRMIT	pWV01 ts/pUC (pBR322)	*aad9*/*aphA3* (for Tn)	Mariner Kan-Tn flanked by MmeI For TnSeq	42
Reporter	pGCP070	pWVO1/NA	*aphA3*/*catA*	TetR inducible sfGFP	100
	pWAR303	pWV01	*ermC*	Luciferase LuxAB	112

[a]Common antibiotic markers used in group A streptococcus mutagenesis include *aad9* (spectinomycin resistance, Sp^r), *catA* (chloramphenicol resistance, Cm^r), *aphA3* (kanamycin resistance, Km^r) and *ermC* (erythromycin resistance, Em^r). Because group A streptococcus remains sensitive to penicillin and is a preferred clinical treatment, the use of β-lactamase for selection during group A streptococcus mutagenesis is both unethical and prohibited.

ple cloning site from pCR2.1 (Invitrogen), allowing for blue/white screening of insertions via α-complementation in *E. coli*. Plasmids pCRS and pCRK are similar to pGCP213, differing primarily in their antibiotic resistance cassettes (25). The modifications to the standard electro-transformation conditions described by Scott's group greatly increase the transformation efficiency of these ts plasmids (23). The main application of the pG+ host-derived ts plasmids has been in the construction of in-frame deletions and allelic replacement to introduce site-specific mutations of chromosomally encoded genes and in the delivery of transposable elements for mutagenesis (see below).

These plasmid-based methodologies can be adapted for functional studies to validate the contribution a given gene confers to a given phenotype via expression in a heterologous host. The heterologous host can be an *S. pyogenes* strain which naturally lacks the phenotype under investigation or can be another streptococcal or enterococcal species. In one of the first examples of this strategy, Scott and Fischetti's group introduced a copy of the gene which encodes the M6.1 protein (*emm6.1*) on a mobilizable plasmid vector, used this to transform a naturally competent streptococcal species, and then mobilized the plasmid by conjugation into an *emm*-deficient *S. pyogenes* host (26). Analysis of the resulting strain demonstrated that the introduction of *emm6.1* converted the host strain from being sensitive to killing in a bactericidal assay to being resistant to killing in the assay. This study was the first formal application of Falkow's molecular Koch's postulates to *S. pyogenes* pathogenesis (26). Additional applications of this approach have included the demonstration that protein F is sufficient to confer a fibronectin-binding phenotype to non-fibronectin-binding *S. pyogenes* and enterococcal hosts (27) and that the *has* operon is sufficient to confer the ability to produce hyaluronate to both acapsular *S. pyogenes* and enterococcal hosts (28).

The second class of commonly used vectors in *S. pyogenes* was based on the low-copy pAMβ1 replicon originally isolated from *Enterococcus faecalis* (29). Unlike pWV01, the pAMβ1 replicon cannot replicate in *E. coli*, so derivatives have been engineered to contain an origin of replication that is proficient in an *E. coli* host (19). The

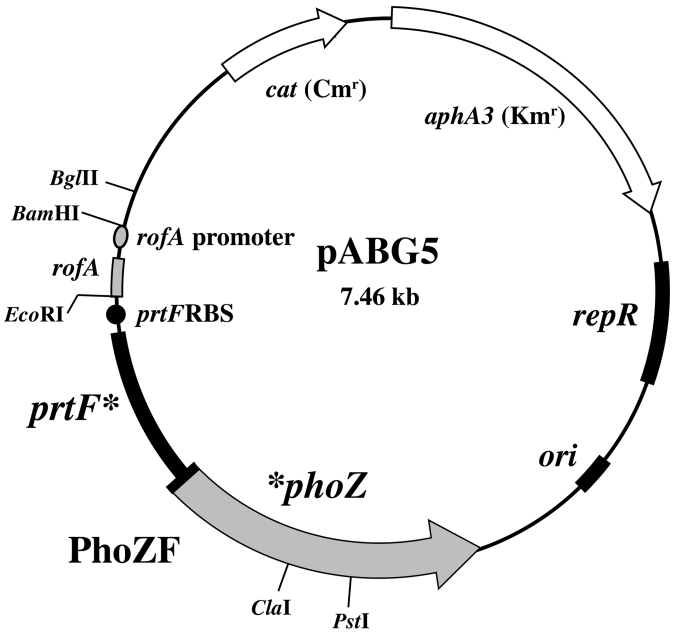

FIGURE 1 A chimeric secreted alkaline phosphatase reporter gene. The plasmid pABG5 is a pWV01-based *E. coli*-streptococcal shuttle vector which encodes resistance to chloramphenicol (*cat*) and kanamycin (*aphA3*). The plasmid contains a promoterless reporter gene formed by the fusion of the N-terminal region of the cell wall-associated protein F (*prtF**) to the enzymatic domain of the enterococcal alkaline phosphatase (**phoZ*). Since the chimeric protein (secreted protein F-PhoZ reporter, or PhoZF) lacks the C-terminal cell wall attachment domain of protein F and the N-terminal lipoprotein tethering domain of PhoZ, it is freely secreted from the cell. The PhoZF chimera retains the enzymatic activity of PhoZ and can be easily quantitated in culture supernatants. Restriction sites for *Bgl*II or *Bam*HI and *Eco*RI can be used to place the promoter of interest in an orientation to direct transcription of *phoZF*, which is then translated using the ribosome-binding site of *prtF* (RBS). Insertion of an open reading frame of interest between the *Eco*RI and either the *Cla*I or *Pst*I site places the open reading frame under the control of the strong *rofA* promoter for ectopic expression and/or complementation analyses.

plasmid pAT28 developed by Courvalin's group (30) is typical of the pAMβ1-derived vectors and contains the high-copy pUC replicon derived from ColE1 for replication in *E. coli* and several other useful features including the multiple cloning site and the *lacZα* reporter gene of pUC18, which allows screening for insertion of cloned DNA by α-complementation in *E. coli* (30). In contrast to the rolling circle replication of pWV01, pAMβ1 replicates via a double-stranded θ-type mechanism which typically leads to greater plasmid stability. Also, as a general rule, pWV01-based vectors transform *S. pyogenes* at a much higher efficiency than do pAMβ1-derived vectors. Because of their different modes of replication, this difference may involve how the restriction system of *S. pyogenes* recognizes the two plasmids. However, it should be noted that very little is currently understood about restriction in *S. pyogenes* and that, in general, it has not posed a significant barrier to the introduction of DNA purified from any number of different *E. coli* K12 hosts. Despite its lower transformation efficiency, the pAMβ1-derived vectors have proven very useful in obtaining expression of cloned genes that were only poorly expressed by pWV01-based vectors (31). A useful derivative of pAMβ1 is pEU7742 (32), which includes a toxin-antitoxin cassette (axe-txe) from an *Enterococcus faecium* plasmid for enhanced plasmid stability.

An alternative to using replicating plasmids for expression of heterologous genes or genes *in trans* is the vector system known as p7INT, which contains an *E. coli* origin of replicon to facilitate manipulation in *E. coli*, along with the integrase gene (*int*) and attachment site (*attP*) from the T12 temperate bacteriophage of *S. pyogenes* (33) (Fig. 2). Following its introduction into *S. pyogenes*, *int* catalyzes the site-specific recombination of the entire circular molecule into the chromosomal phage T12 attachment site (*attB*) located at the 3′ end of the transfer-messenger RNA gene (33) by a mechanism resembling integration of lambda phage (34). Because the plasmid lacks the phage excisionase, integration is highly stable. Derivatives of p7INT have been used to study Rgg-dependent gene regulation through use of a *luxAB* luciferase-based reporter system (35). It should be noted that integration of p7INT is strain specific, because two alleles of transfer-messenger RNA (tmRNA) exist among the published sequences, which are distinguished by a single nucleotide polymorphism in *attB*. Common strains, including NZ131 and HSC5, have an *attB* site compatible for p7INT integration, while others, including JRS4, have an incompatible site. However, p7INT can be modified by a single base change, generating p7INT*, which can integrate into the alternative *attB* site (G.C. Port and M.G. Caparon, unpublished) (Fig. 2).

DIRECTED MUTAGENESIS

The techniques developed for directed mutagenesis (reverse genetics) in *S. pyogenes* have proven invaluable for the analysis of the contributions that specific genes make to pathogenesis. In this approach, a defined mutation is constructed to inactivate, or in some cases modify, a specific gene to construct an isogenic mutant strain for analysis. These techniques have become particularly useful with the availability of the complete or draft genome

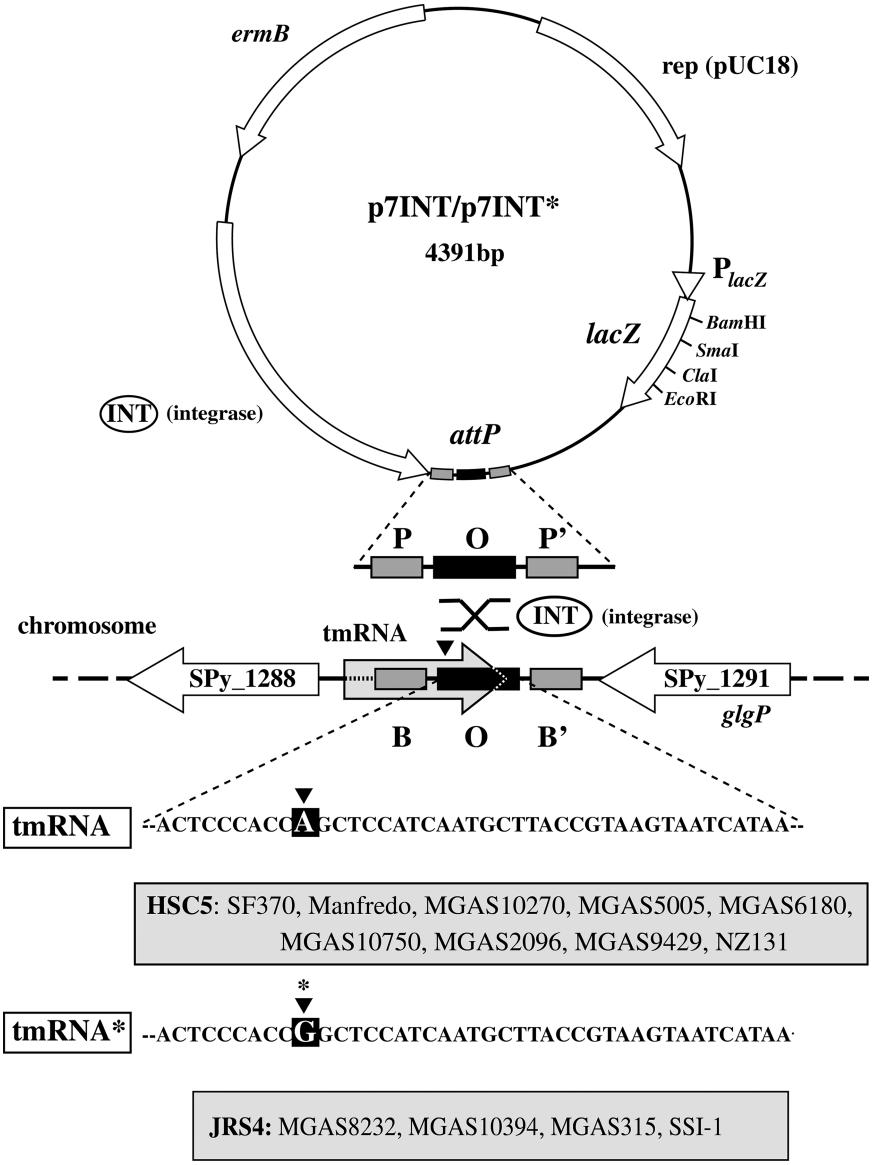

FIGURE 2 p7INT and p7INT*. The plasmids p7INT and p7INT* are pUC18-based *E. coli* vectors that encode resistance to erythromycin (*erm*). Restriction sites for *Bam*HI, *ClaI*, *SmaI*, and *Eco*RI within *lacZ* can be used to insert any promoter and gene of interest for ectopic expression and/or complementation analyses. The plasmids also contain a phage integrase (INT) and an *attP* sequence which targets integration into the *S. pyogenes* chromosome at the 3′ end of the tmRNA locus within the 96-bp "O" site of POP′ and BOB′ of *attP* and *attB*, respectively. With respect to O sites, *S. pyogenes* has been found to contain one of two alleles with either an A or G (see text). Plasmid p7INT contains an A, whereas p7INT* has been mutated to contain a G, and the appropriate plasmid must be used depending on the allele found within the *S. pyogenes* strain of interest. Following introduction of the appropriate plasmid into *S. pyogenes*, selection for the antibiotic-resistant determinant of the plasmid (*erm*) selects for chromosomes in which the plasmid has integrated into the chromosome at the *attB* site. The distribution of tmRNA and tmRNA* between a selection of commonly used strains is shown, with GenBank accession numbers as follows: HSC5, CP006366.1; SF370, AE004092.2; Manfredo, AM295007.1; MGAS10270, CP000260.1; MGAS5005, CP000017.2; MGAS6180, CP000056.1; MGAS10750, CP000262.1; MGAS2096, CP000261.1; MGAS9429, CP000259.1; NZ131, CP000829.1; JRS4, CP011414.1; MGAS8232, AE009949.1; MGAS10394, CP000003.1; MAGS315, AE014074.1; SSI-1, BA000034.2.

sequences of *S. pyogenes* (3). In the following, we consider methods to alter or inactivate a targeted gene on the *S. pyogenes* chromosome using (i) allelic replacement using linear DNA, (ii) directed insertional integration, and (iii) construction of in-frame deletion alleles.

Allelic Replacement Using Linear DNA

This method of mutagenesis is straightforward and involves first cloning the gene of interest into any suitable plasmid vector in *E. coli*. With the available genome information, this has become a simple task of analyzing the sequence and designing suitable primers for PCR amplification, followed by insertion of the amplification product into an *E. coli* cloning vector. As a general rule, *S. pyogenes* DNA is most stable when cloned on low-copy vectors in *E. coli*, and several vectors have been used with great success. The standard PCR methodology is then used to insert a selectable marker into the open reading frame to inactivate it. The markers that have been most successful are those derived from Gram-positive organisms that can also be used for selection in *E. coli* and include *ermC* (resistance to erythromycin; GenBank accession no. NG_047805.1), *aphA3* (resistance to kanamycin; GenBank accession no. V01547.1), *tetM* (resistance to tetracycline; GenBank accession no. X92947.2), *aad9* (resistance to spectinomycin; GenBank accession no. M69221.1), and *catA* (resistance to chloramphenicol; GenBank accession no. NG_047564.1). It should be noted that because resistance to β-lactam antibiotics does not naturally exist among clinical *S. pyogenes* isolates, for ethical reasons, genes encoding any type of β-lactamase must not be used for this application. Strong polar mutations can be introduced through the use of dedicated insertional mutagenesis cassettes, such as the Omega interposon (ΩKm-2), which contains strong transcription and translation termination signals (36). Following its construction, the insertionally inactivated allele is converted to a linear molecule by PCR amplification or restriction digestion of the backbone plasmid. The resulting linear molecule is then used to transform the target *S. pyogenes* host with selection for the inserted resistance marker. Transformation of linear DNA is best conducted using highly purified salt-free DNA (for example, see 37). Because the introduced molecule is linear, preservation of the circular chromosomal structure requires that all resistant transformants arise by two homologous recombination events flanking each side of the inserted resistance marker. The end result is the exchange of the wild-type allele for the inactivated allele (Fig. 3). While this type of recombination does not occur at high frequency, the method is usually successful because the frequency of nonhomologous recombination is typically extremely low. Insertion into the expected locus is confirmed using PCR-based analyses and/or genome sequencing, and the resulting mutant can be subjected to functional studies. However, in interpreting the resulting functional data, it should be kept in mind that this method of mutagenesis generates strong polar mutations. Thus, expression of distal genes should be examined using quantitative real-time reverse transcriptase PCR (qRT-PCR) or other methods suitable for analysis of gene expression.

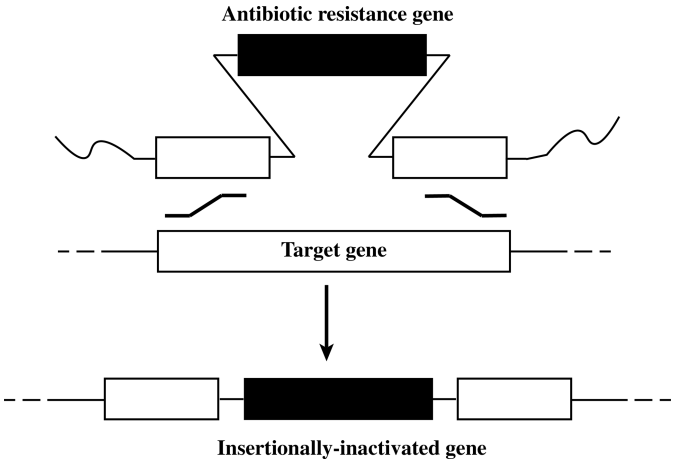

FIGURE 3 Strategy for allelic replacement mutagenesis using linear DNA. Segments of DNA homologous to the 5′ and 3′ ends of the target gene are placed flanking a gene encoding a selectable antibiotic resistance gene, typically on a standard plasmid vector using *E. coli*-based molecular cloning technology. The resulting *E. coli* plasmid vector is converted to a linear molecule by PCR or by digestion using restriction enzymes with sites outside of the cloned streptococcal gene (vector DNA is represented by the curved lines) and introduced into *S. pyogenes* with selection for resistance to the antibiotic encoded by the introduced resistance gene. Recombination between the two homologous sequences (indicated by the lines between the introduced DNA and chromosome) results in the replacement of the insertionally inactivated allele for the native chromosomal allele (shown below the arrow).

A variation of this technique can be used to subject a large cloned region to rapid high-resolution insertional mutagenesis, which can be useful for analyses of complex phenotypes and pathways whose genes are often clustered together on the chromosome. This method uses a version of the *E. coli* transposon mini-γδ (mγδ-200) that has been modified to contain a kanamycin-resistance gene that can be selected for in both *E. coli* and *S. pyogenes*. Called "shuttle mutagenesis," this technique uses the ability of m-γδ to easily and efficiently generate a large number of random transposon insertions into a segment of DNA

cloned on a plasmid to obtain a series of mutations along the entire length of the cloned streptococcal DNA. These sets of nested insertions are then crossed into the streptococcal chromosome as described above. This method has been successfully used to identify regulatory genes linked to a target gene of interest (38). For a detailed description of this technique, consult Hanski et al. (39).

Directed Insertional Inactivation

A commonly used technique for directed mutagenesis employs mutations constructed as the result of a single homologous recombination event. For this technique, an internal segment of the target gene lacking the 5′ and 3′ ends is cloned into a Gram-negative restricted plasmid. A selectable marker for *S. pyogenes* is also included on the plasmid located adjacent to the cloned segment, but this marker should not be a β-lactamase (see above). As a general rule, the larger the fragment that is amplified, the greater the frequency at which allelic exchange occurs, although this method has been successful with fragments in the 250- to 500-bp range. The resulting plasmid is purified from *E. coli* and used to transform *S. pyogenes* with selection for resistance to the introduced marker. Because the commonly used *E. coli* replicons do not replicate in *S. pyogenes*, resistant transformants most frequently arise as a result of a single homologous recombination event between the internal segment of the gene introduced on the plasmid and the identical sequence on the streptococcal chromosome. The resulting chromosomal structure is generated by integration of the entire plasmid and consists of a partial duplication of the target gene, which now flanks the integrated plasmid sequences (Fig. 4A). As a consequence of using only an internal fragment of the target gene, one of the duplicated copies lacks its 3′ end and the other lacks its 5′ end such that both partial copies should be inactive. Thus, the end result is the directed insertional inactivation of the target gene. A number of modified *E. coli* plasmids have been utilized for this mutagenesis technique, often by replacing a β-lactamase ampicillin cassette for an appropriate selection marker. These include pCIV2 (40) and pSPC18 (41), which contain a kanamycin and spectinomycin resistance cassette, respectively.

The principal advantage to this method is that single recombination events are relatively quick and easy to generate because they require a single selection step rather than the multiple passages and screening required for construction of in-frame deletions using ts shuttle vectors (see below). This approach is also simple because it requires the cloning of only a single segment of the chromosome and can work with fragments as small as 0.25 kb. However, there are several disadvantages to the method. These include the possibility that the mutations are not necessarily stable because a second recombination event between the duplicated gene segments can result in excision of the plasmid vector and regeneration of a wild-type structure. Also, because the incoming plasmid does not replicate, it has only a limited period of time with which to undergo recombination for insertion into the chromosome. This latter limitation was addressed by Le Breton et al., who devised a two-plasmid system whereby a suicide plasmid (pSinS) lacking the pWVO1 RepA is complemented by a helper plasmid (pHlpK) that provides a ts RepA. The two plasmids are used to transform *S. pyogenes* at a permissive temperature, and a subsequent shift to a higher temperature prevents plasmid replication, selecting for plasmid integration (42). A final limitation to discuss is that the mutations generated by this technique are strongly polar. However, it is possible to modify the technique to directly assess the impact of polarity on the phenotype under investigation. In this case, an additional control strain is constructed, where instead of cloning a segment internal to the target gene, a segment is cloned that has its 5′ end anchored internal to the gene, but its 3′ end is anchored at a location distal to the 3′ end of the coding region of the gene. Integration of this construction into the target gene in a wild-type strain also results in a partial duplication. However, because the cloned segment overlaps the 3′ end of the gene, the first copy of the gene is regenerated and the integrated vector is located adjacent to this intact copy such that it is still polar on expression of any distal genes (Fig. 4B). If the phenotype under analysis is the result of a polar effect and the loss of expression of a distal gene, then this control strain should also demonstrate the mutant phenotype, even though it has an intact and functional copy of the target gene. On the other hand, an unaltered wild-type phenotype in the control strain indicates that the mutant phenotype is solely the result of insertional inactivation of the target gene (for example, see 43).

An additional modification of this method can be used to map the promoter and *cis*-acting control regions of a target operon. This is based on the technique developed by Moldover and colleagues (44) for analysis of promoters in *Bacillus subtilis* and is similar in concept to the method described above for placing a polar insertion downstream of a target gene. However, in this technique, the segment of DNA cloned into the integrational vector is anchored within the coding region at its 3′ end but is anchored upstream of the start of the gene at its 5′ end. Integration of this construct into the target locus again

generates duplication. However, since the 3′ end of the cloned segment ends within the gene, the first copy is truncated and is inactive. This is followed by the integrated vector and then by the second intact copy of the target gene, which is now preceded exactly by the 5′ flanking region cloned into the vector. If this 5′ flanking region includes the promoter and other *cis*-acting control regions, the target gene will be expressed (Fig. 4C-1). If this segment lacks these elements, the target gene will not be expressed (Fig. 4C-2). By using a nested set of insertions containing different lengths of the upstream control region, it is possible to map the sequences required for expression and regulation of the target gene with a high degree of precision. This method has been used to examine regulation of *mga*, the regulator of the genes which encode the M proteins and the C5a peptidase (40).

In-Frame Deletion

Despite the increase in time and effort, the construction of in-frame deletions has become the standard for directed mutagenesis because they are stable and nonpolar. In-frame deletions have been constructed in the genes for many *S. pyogenes* virulence factors, including the C5a peptidase (16), the M protein, the cysteine proteinase (45), the *has* operon (46), and streptolysin O (47). Most in-frame deletion mutations were constructed by the method pioneered in the Cleary lab (16) that utilizes a derivative of the ts plasmid pG+host4. In this method, the gene of interest is cloned into the vector in *E. coli*. Standard molecular techniques are then used to delete a large central region of the gene, but at the same time preserve its reading frame. Alternatively, a set of four primers are designed to perform splicing by overlap extension (48) to directly generate the deletion construct and insert it into the shuttle plasmid, all via PCR and without the use of restriction digestion and ligation (49). The resulting construct is introduced into an *S. pyogenes* host at a temperature that is permissive for replication of the plasmid. The culture is then shifted to the nonpermissive temperature while maintaining selection for the resistance determinant of the plasmid. This selects for chromosomes into which the nonreplicating plasmid has been inserted by homologous recombination between the in-frame deletion allele and the resident wild-type allele. At this stage, the chromosome contains both the wild-type and deletion alleles. At some frequency, a second homologous recombination occurs which results in the excision of the integrated plasmid. Depending on the recombination junctions, either the wild-type or the deletion allele remains behind in the chromosome (Fig. 5). Shifting the culture back to the permissive temperature enriches for chromosomes from which the plasmid has been excised, since replication of the integrated plasmid creates a second origin of replication for the chromosome, which is usually deleterious to growth. The culture is once again shifted to the nonpermissive temperature, but this time in the absence of selection for the plasmid. This enriches for segregants, which have lost the excised and now nonreplicating plasmid. These segregants are plated as single colonies, which are then examined by streaking each colony onto agar plates in the presence or absence of antibiotic selection to identify clones which have lost the plasmid and thus are sensitive to the antibiotic used for selection. These are typically subjected to further screening by PCR and/or genome sequencing to identify chromosomes which contain the in-frame deletion allele. For a detailed description of this technique see Le Breton and McIver (12).

Counterselection Markers

The requirement for multiple passages at nonpermissive and permissive temperatures can make this method of in-frame deletion time-consuming. However, counterselection strategies that select against the retention of the plasmid can be used to expedite the recovery of chromosomes that have undergone plasmid excision and loss (50). One such strategy involves phenylalanyl tRNA synthetase (alpha-subunit) encoded by *pheS*. A single amino acid substitution in PheS (PheS*) confers relaxed substrate specificity and allows for the misincorporation of a toxic derivative of phenylalanine, 4-chloro-phenylalanine (4CP), into proteins. Here, a plasmid containing *pheS** is integrated into the chromosome. Subsequent culture in the presence of 4CP results in the death of cells that retain the *pheS**-containing plasmid, while those that have excised the plasmid survive (51). The addition of a second amino acid substitution in PheS* (defined here as PheS**) (52) confers even greater misincorporation of 4CP and higher counterselection efficiency and has been successfully used in *Streptococcus mutans* (53) and in *S. pyogenes* (Port and Caparon, unpublished) (Fig. 6). An alternative counterselection strategy involves introduction of the *rpsL* marker into the ts vector (54). This gene encodes ribosomal protein S12, and point mutations in *rpsL* which render the ribosome resistant to the antibiotic streptomycin arise spontaneously in *S. pyogenes* at a frequency of $\sim 10^{-9}$ cell^{-1} generation^{-1} (26). However, *in trans*, the wild-type allele is dominant, so when a ts vector containing wild-type *rpsL* is integrated into the chromosome of a streptomycin-resistant strain, the resulting partial diploid will be sensitive to streptomycin.

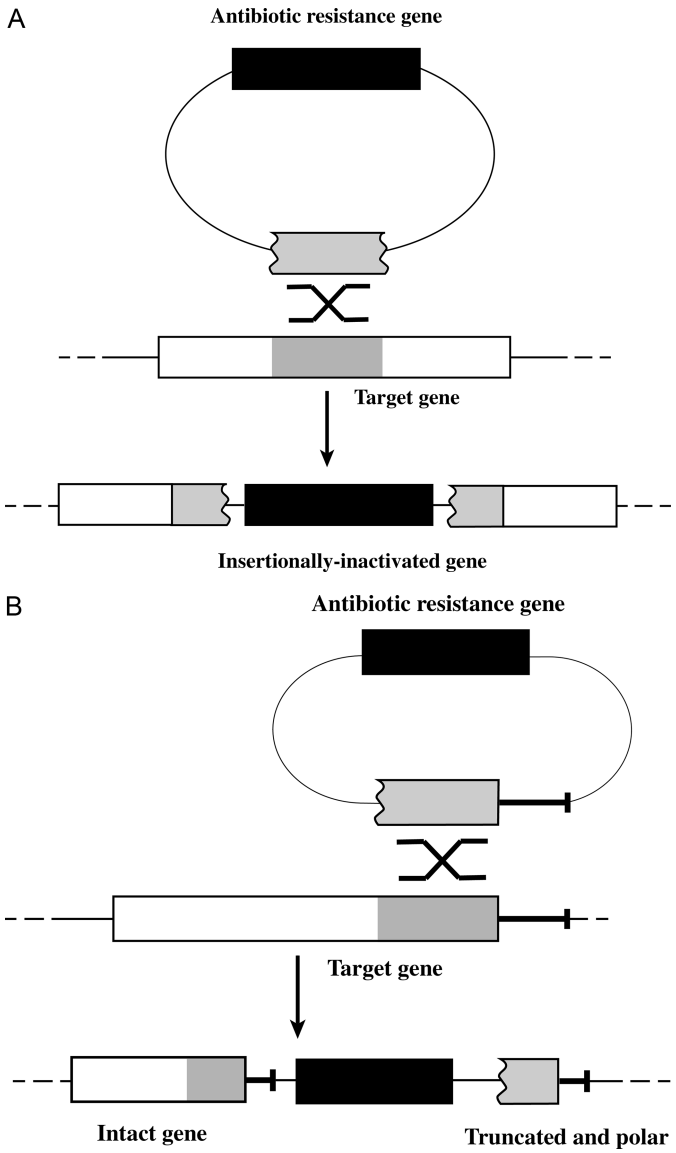

FIGURE 4 Directed insertional mutagenesis of targeted genes using circular DNA. **(A)** Insertional inactivation of the target gene. A DNA segment internal to the target gene (shown by the gray box enclosed by wavy lines) is cloned onto an *E. coli* plasmid which cannot replicate in *S. pyogenes*. The plasmid is introduced into an *S. pyogenes* host as a circular molecule (top of figure) with selection for a resistance marker on the plasmid. A single homologous recombination event between the chromosome and the circular molecule (shown by the "X") results in the integration of the plasmid into the chromosome, a partial duplication of the gene (the two regions shaded in gray) in which neither of the two copies are complete, resulting in gene inactivation. **(B)** Generation of a polar insertion 3′ to the target gene. If the segment of DNA cloned on the integrational plasmid includes the 3′ terminus and sequences downstream of the target gene (shown by the regions shaded in gray and the thick bar, respectively), homologous recombination (shown by the "X") results in a partial gene duplication in which the 5′ copy is intact but the 3′ copy is truncated, producing a polar effect on any downstream gene(s) which may share a promoter in common with the target gene. When compared with an insertion mutation that inactivates the target gene (Fig. 1 or Fig. 2), this strategy tests whether a phenotype is due to the loss of target gene function or to a polar effect on a distal gene. **(C)** Mapping the *cis*-acting control regions of the target gene. The technique is modified by constructing several plasmids representing a nested set of the chromosomal region extending various distances 5′ of the target gene. The plasmids are integrated into the target locus as described above, and the end-products are a partial duplication of the target gene in which the proximal copy is truncated and inactive. If the duplicated region includes the *cis*-acting control regions (represented by the broken arrow and the closed circle labeled "P"), the distal copy of the target gene will be expressed (scenario 1). In contrast, if the cloned segment does not include the *cis*-acting control regions, then the distal copy will not be expressed (scenario 2).

Figure 4 continues on next page

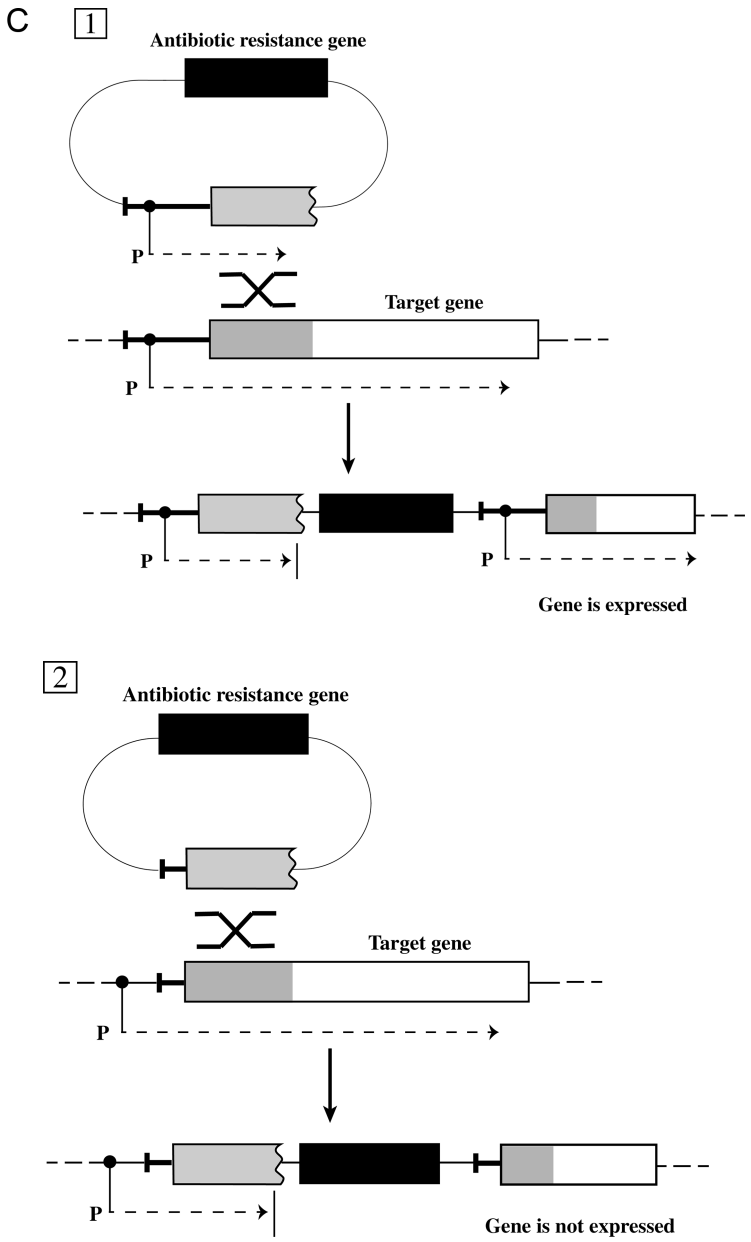

FIGURE 4 *(continued)*

Thus, the inclusion of streptomycin during the final passage at the nonpermissive temperature directly selects for chromosomes from which the integrated plasmid has excised and segregated. This method has been used to introduce mutant alleles of the virulence regulator Mga into the chromosome for functional analyses (54).

In-Frame Deletion Using CRISPR/Cas9

A drawback to the standard in-frame deletion method mentioned above is that it can be time-consuming, usually taking more than a week for the multiple passages that are required to enrich for chromosomes that contain the deletion allele (Fig. 5). This time period can be reduced through the application of a system using clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9). CRISPR/Cas9 was originally described in *S. pyogenes*, where it functions to confer an adaptive immunity against invading genetic elements, especially bacteriophages (55, 56). It does this by introducing double strand breaks at specific sequences that the Cas9 enzyme is programmed to recognize by incorporation of a small targeting RNA molecule (CRISPR

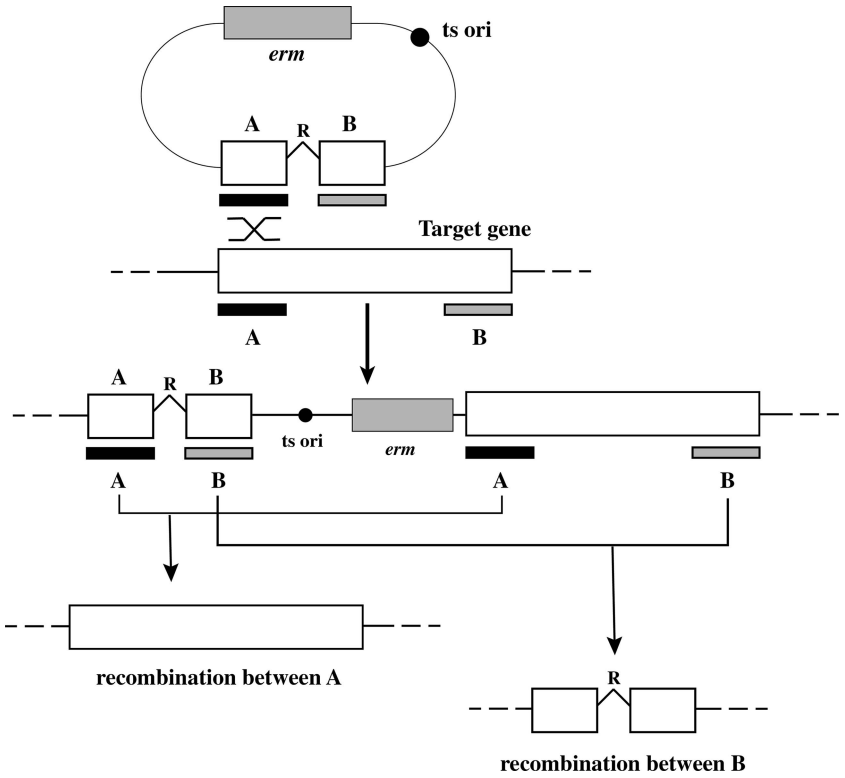

FIGURE 5 Construction of an in-frame deletion. Standard PCR-based methods are used to generate a deletion of the internal region of a copy of the target gene that has been cloned on an *E. coli*-streptococcal shuttle vector that is temperature-sensitive for replication (indicated by "ts ori"). The deletion is constructed so as to maintain the reading frame of the gene (represented as the bent line labeled "R," connecting the 5′ region labeled "A" and the 3′ region labeled "B"). Following its introduction into *S. pyogenes*, growth at a temperature that is nonpermissive for replication of the plasmid with selection for the antibiotic-resistant determinant of the plasmid (*erm*) selects for chromosomes in which the plasmid has integrated by homologous recombination (indicated by the "X"). The two regions of homology flanking the deletion are represented by the solid and gray bars labeled "A" and "B." Recombination between the A bars or, alternatively, between the B bars (shown by the brackets and arrows) results in excision of the plasmid and either the restoration of the wild-type structure or allelic exchange for the deletion allele (these products are illustrated below the second set of arrows). Growth at a temperature permissive for replication of the plasmid enriches for excisant chromosomes, presumably because if integrated, the plasmid origin of replication becomes active, resulting in misregulation of chromosomal replication. Presence of the wild-type or deletion allele in any isolate is easily determined by assay for the unique restriction site engineered into the deletion allele (indicated by the "R" above the bent line) or by PCR with primers flanking the region of interest.

RNA [crRNA]) that is encoded in the CRISPR array (55, 56).

To cleave a specific DNA site, the Cas9 nuclease associates with both a crRNA that has a short sequence complementary to the DNA target sequence (known as the proto-spacer) and with a separate trans-activating crRNA (trRNA, sometimes called tracrRNA) (57). The trRNA is partially complementary to the crRNA and is required for crRNA maturation from a primary transcript encoding pre-crRNAs. The crRNA and trRNA form a based-paired structure that is incorporated into Cas9 and targets the complex to a proto-spacer region located adjacent to a short motif known as the protospacer-adjacent motif, which for *S. pyogenes* Cas9 is NGG. Jinek et al. simplified the system by merging trRNA and crRNA into a single synthetic molecule known as the single guide RNA (sgRNA) (57). Due to its simplicity, requiring only Cas9 and sgRNA, the type II CRISPR/Cas9 system from *S. pyogenes* has found wide application for genome editing in many organisms, including higher eukaryotes

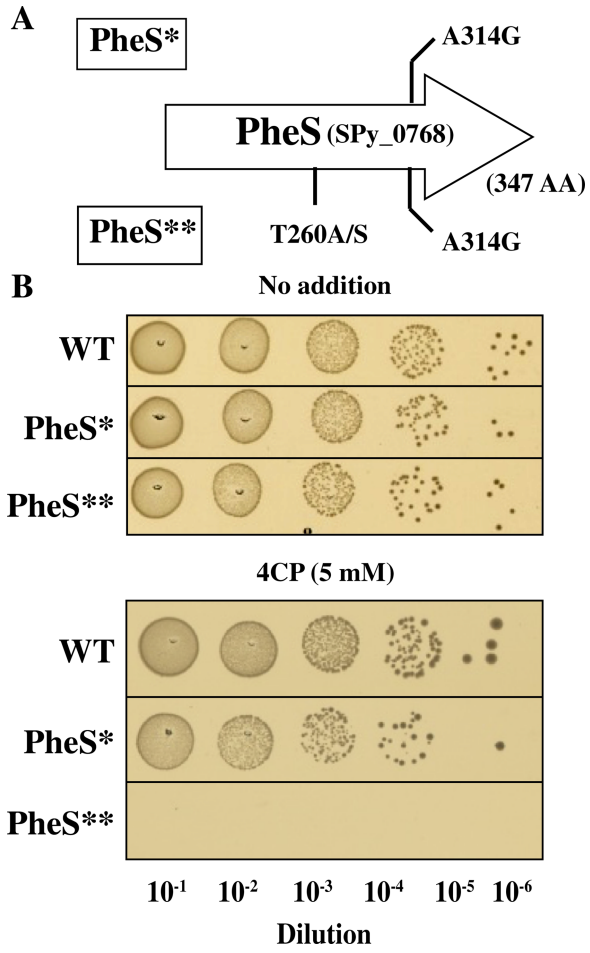

FIGURE 6 PheS**, a counterselectable marker. The phenylalanyl tRNA synthetase (alpha subunit) PheS catalyzes the attachment of phenylalanine onto its cognate transfer RNA. **(A)** Introduction of a single amino acid substitution (PheS*, A314G) confers relaxed substrate specificity and allows for the misincorporation of a toxic derivative of phenylalanine, 4-chloro-phenylalanine (4CP), into proteins. Addition of a second substitution at position 260 (PheS**, either T260A or T260S in combination with A314G) facilitates increased misincorporation of 4CP. **(B)** *S. pyogenes* containing a multicopy plasmid (pABG5) either as an empty vector (WT) or engineered to express PheS*, or PheS** was serial diluted and plated onto solid media in the absence (top panel) or in the presence (bottom panel) of 5 mM 4CP. Strains expressing PheS* show moderate sensitivity to 4CP toxicity, whereas strains expressing PheS** show enhanced sensitivity.

(58). A derivative, "dead" Cas9 (dCas9), modified by site-specific mutations that inactivate its two nuclease activities, can be used in a variety of approaches for silencing expression of specific genes (the CRISPRi method) (56, 59).

Cas9 can also be adapted for the efficient construction of in-frame deletions in *S. pyogenes* by providing a direct counterselection against strains that retain the wild-type allele. This method requires modification of the standard ts plasmid to contain Cas9 and an sgRNA in addition to the in-frame deletion allele (K.H. Cho, unpublished). The sgRNA should include a short ~20-nucleotide sequence complementary to a sequence in the targeted gene that will be eliminated by the deletion, which is adjacent to a suitable protospacer-adjacent motif. When the plasmid is introduced into *S. pyogenes* with the wild-type gene allele, only the strain with the in-frame deletion allele in its chromosome created by double-crossover survives, because it has eliminated the protospacer sequence. In contrast, any strain retaining the wild-type allele is killed due to the introduction of a double strand break of its chromosome by the Cas9 nuclease. This process is highly efficient, because unlike most eukaryotic organisms that can repair a double strand DNA break through nonhomologous end joining and homology-directed recombination (60), many prokaryotes use only homology-directed recombination to repair the double strand break. Thus, if no homologous DNA is available for repair, the double strand DNA break can be lethal (61). This method has been successfully used to create in-frame deletions in *Staphylococcus aureus*, *Streptococcus pneumoniae*, and *E. coli* (62, 63). It is anticipated that this method will be even more successful in *S. pyogenes*, because more than half of the *S. pyogenes* genomes (14 out of 20 completely sequenced and annotated strains available in GenBank at the time of this writing) already possess *cas9*. For these strains, just providing an sgRNA may be sufficient for selection of in-frame deletion mutants.

Considerations for Directed Mutagenesis

The multiple *in vitro* passages required for directed mutagenesis techniques have a tendency to enrich for spontaneous mutants with reduced expression of virulence factors (64), including the M protein and the hyaluronic acid capsule (65). It is therefore important to isolate and phenotypically test multiple independently derived mutants to confirm that they have a consistent phenotype, as well as to genetically complement any mutants to confirm that the targeted mutation is the cause of the observed phenotype. If the gene targeted for mutagenesis lies upstream of an essential gene, it may be difficult to select for the initial single crossover due to a polar effect. To counter this point, a larger DNA fragment encompassing the entire downstream gene can be utilized to ensure continued gene expression following the initial crossover (66). Finally, some mutants are less fit and have a slower growth rate than the wild type. Therefore, any deletion mutants that are generated are quickly outcom-

peted by their wild-type counterparts, making them difficult to recover following *in vitro* passage. In this case, an allelic replacement mutant can be made whereby a gene of interest is replaced by an antibiotic cassette, which allows for direct selection of the mutated chromosome (66).

TRANSPOSON MUTAGENESIS

The techniques described above have been extremely useful for the analysis of virulence in *S. pyogenes*. However, they require a detailed pre-existing knowledge of the target gene and thus are more suited for reverse genetics approaches applied to testing hypotheses which address the functions of previously identified and characterized genes or genes identified from whole-genome sequence information whose function is suggested by homology with a characterized gene. This approach is not well suited for a forward genetics approach for the identification of novel genes that contribute to virulence. This is most successfully done using a traditional unbiased genetic approach, in which a virulence phenotype is established and random mutations are generated to identify genes which influence this phenotype. In *S. pyogenes*, the most successful method of random mutagenesis has used the insertion of a transposable element at multiple loci in the chromosome. This section considers the various types of transposons that have been successfully used to identify novel genes in *S. pyogenes*, with particular focus on recently adapted transposon systems that allow for the characterization of high-diversity libraries, which include hundreds (signature tagged mutagenesis) or even thousands (TnSeq) of mutants during infection.

Tn*916*

The first transposable element used to identify novel virulence genes in *S. pyogenes* was Tn*916* (Fig. 7A). This element is the prototype of the family of transposons discovered by Clewell's group (29), which are now known as the conjugative transposons. They have the remarkable property that they are self-transmissible and transpose from a locus in a donor chromosome to a different locus in a recipient chromosome when these chromosomes are located in different cells. The conjugal transfer event requires cell-cell contact and can occur between very distantly related species and even between Gram-positive and Gram-negative hosts. This biology of these unique elements has been extensively investigated and is the subject of several excellent reviews (for example, 67) and thus will only be briefly mentioned here.

The advantage to using Tn*916* is that it can be applied to probably any isolate of *S. pyogenes*, including those which may be difficult to transform. The disadvantages are that Tn*916* is a large element for a transposon (over 18 kb in size), does transpose into preferred sites, and frequently generates chromosomes with multiple insertions. The transposon has served as a useful mutagenesis agent and has been used to identify novel genes involved in regulation (e.g., *mga* [7]), expression of streptolysin S (68–70), and production of the capsule (71, 72), among others.

Tn*917*

This transposable element was also originally described by Clewell's group in *E. faecalis* (73), highly engineered by Youngman and colleagues, and applied with great success to the mutagenesis of *Listeria monocytogenes* and *B. subtilis*. One characteristic of this family is that its members contain fairly large terminal repeats, although they are much smaller in size than the conjugative transposons. They produce stable insertions, generate a 5-bp target site duplication upon insertion, and transpose at high frequency and with a high degree of randomness. Their movement involves a replicative pathway, which requires that donor and recipient molecules undergo replication, and proceeds through the cointegration of donor and host molecules which contain two directly repeated copies of the transposon at the fusion junctions. Resolution of the cointegrate occurs via site-specific recombination catalyzed by a transposon-encoded enzyme called resolvase that recognizes a specific site in each copy of the transposon (the *res* site). The end result is that a single copy of the transposon is located at a random site in the recipient molecule. Scott's group (74) developed a mutagenesis vector (pJRS290) (75) that is based on pG+host4 and contains a highly engineered derivative of Tn*917* developed by Youngman and coworkers (76) (Tn*917*-LTV3) (Fig. 7B). This derivative of Tn*917* has several useful features, including an *E. coli* plasmid origin of replication that facilitates the direct cloning of chromosomal DNA adjacent to the inserted transposon. Using this vector, Tn*917*-LTV3 was found to transpose efficiently and with a high degree of randomness in *S. pyogenes*.

Tn*4001*

The transposable element Tn*4001* was originally isolated as an agent of transmissible gentamycin-resistance in *S. aureus* (77). It is a composite-type transposon and a member of the Tn*5* family. These transposable elements share a common structure in which directly repeated copies of an insertion sequence flank an antibiotic

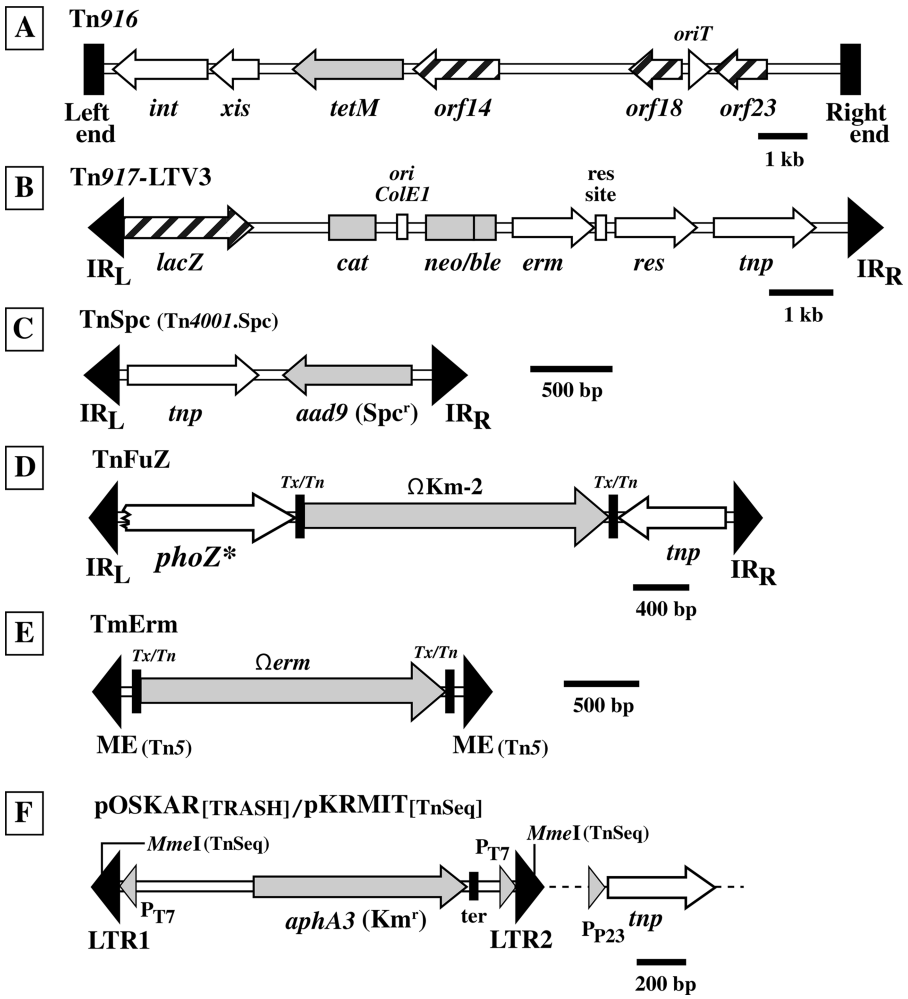

FIGURE 7 Transposon mutagenesis of *S. pyogenes*. The several transposons that have been used for mutagenesis of *S. pyogenes* are shown. Antibiotic resistance genes are represented by the gray bars, transposon ends and/or terminal inverted repeats are shown in black, genes that are essential for transposition are shown by white bars, and genes that are nonessential for transposition are represented by the stripped bars. **(A)** Tn*916* (GenBank accession no. U09422) is the prototype conjugative transposon and contains at least 24 different open reading frames. Of these, one is required for resistance to tetracycline, two are essential for transposition, and the rest are likely involved in conjugal transfer. **(B)** Tn*917*-LTV3 is a highly engineered derivative of the Tn*3*-like transposon Tn*917*. This element transposes via a replicative mechanism and has been modified to include a promoterless *lacZ* reporter gene to generate random transcriptional fusions and an *E. coli* ColE1 plasmid origin of replication to facilitate the cloning and analysis of inactivated loci. **(C)** TnSpc is a derivative of the Tn*5*-like transposon Tn*4001*, which transposes via a cut and paste mechanism and consists of the left and right inverted repeats and transposase of IS*256* and a spectinomycin resistance gene. **(D)** TnFuZ introduces the gene for the *E. faecalis* alkaline phosphatase (*phoZ*) altered by removal of the region encoding its signal sequence (broken line at the 5′ end of *phoZ**). TnFuZ acts as an "export signal sequence trap." Insertions into genes that encode a protein export sequence promote the secretion of PhoZ* enzymatic activity, which can be detected by a number of high-throughput assays for alkaline phosphatase activity. **(E)** TmErm is an engineered element that contains the right and left mosaic ends (ME) of Tn*5* flanking an erythromycin resistance marker (Ω*erm*) that includes strong transcription and translation stop sites (*Tx/Tn*). To apply transposome mutagenesis to *S. pyogenes*, the DNA fragment containing TmErm is purified, such that the resulting preparation does not include the ampicillin resistance determinant contained on the pMOD-2::Ω*erm* vector backbone. The purified DNA fragment containing TmErm is then reacted with transposase (EZ::TN, Epicentre Technologies) *in vitro* to form the transposome that is then used to transform *S. pyogenes* using electroporation. **(F)** pOSKAR and pKRMIT are transposon systems incorporated into the ts pWV01-based *E. coli*-streptococcal shuttle vector that utilize the highly active *mariner Himar1* transposable element (MarC9) under control of the P23 promoter from *Lactococcus lactis*. Both transposon systems include a kanamycin resistance gene (*aphA3*) flanked by long terminal repeats (LTR) whose promiscuous target insertion site requires only a TA dinucleotide sequence, leading to a highly random library of insertions. pOSKAR can be utilized for TraSH analysis through the use of outward-facing T7 promoters, while pKRMIT also includes an outward-cutting *MmeI* site for use with TnSeq analysis.

resistance gene. Each insertion sequence itself is bounded by short inverted repeats that flank a transposase gene, which is the only transposon-encoded gene required for movement. The individual insertion sequences themselves are usually capable of transposition independent of the entire element. For Tn*4001*, two copies of the insertion sequence IS*256* flank a central gentamycin-resistance gene. The transposition pathway for this class of elements involves a "cut and paste" mechanism catalyzed by the transposase, which recognizes as its substrates the short inverted repeats of the insertion sequences. Studies have shown that Tn*4001* has a broad host range, including staphylococci, oral streptococci, and mycoplasmas, and that it chooses its targets for insertion with a high degree of randomness. Since the cut and paste pathway does not require replication of the donor molecule, the element can easily be delivered into the host of choice via a nonreplicating suicide vector. However, a number of problems have limited its use. The most significant of these is that transposition of the entire element occurs at a somewhat low frequency relative to the frequency of transposition of the individual insertion sequences. As a result, a chromosome with a Tn*4001* will frequently also have multiple copies of IS*256* inserted at other loci. To address this problem, derivatives of Tn*4001* have been constructed that essentially contain only the inverted repeats of IS*256* oriented to flank the gene for transposase and a selectable marker (78). The resulting elements are relatively small in size (~2 kb) and transpose at a high frequency characteristic of an individual copy of IS*256*. The organization of the element also prevents independent transposition of the insertion sequence and ensures that the resulting population of insertions is homogeneous, although the resulting mutants may have insertions at two loci. A spectinomycin-resistant version of this transposon (called TnSpc for simplicity) (Fig. 7C) was used to identify novel genes required for expression of the SpeB cysteine protease of *S. pyogenes*, including a transcriptional regulator (*ropB*) and a chaperone (*ropA*) (78). Additional variants include the introduction of an erythromycin resistance gene and 1 of 12 unique 34-bp signature tags for application in signature-tagged mutagenesis studies. In this method, 12 independent pools of mutants are generated. Then, one mutant is taken from each of these pools, and they are combined to generate an inoculum with 12-fold complexity that is used as the input into an assay of virulence. The output pool is then analyzed by a PCR assay to conduct a census of the recovered sequence tags, which is then compared with the input pool to determine the relative fitness of individual mutants. Mutants with reduced fitness are putative virulence genes, and this method has been successfully applied to the identification of genes that promote pathogenesis in a zebrafish model of *S. pyogenes* infection (79, 80).

Tn*FuZ*

Since the terminal inverted repeats of IS*256* are quite short (26 bp), several specialized derivatives of the transposon have been developed that place a reporter gene adjacent to the IS*256* end. The Tn*FuZ* (fusion to *phoZ*) reporter places a version of the *E. faecalis phoZ* alkaline phosphatase adjacent to the transposon inverted repeat (Fig. 7D). This reporter gene has been modified by removal of its endogenous export signal sequence to construct an "export signal sequence trap." In this method, insertion of the element into an open reading frame can result in the formation of a translational fusion with PhoZ (the use of PhoZ as a reporter gene is discussed below). If the donor open reading frame contains a secretion export signal, the chimeric PhoZ protein will be secreted. Since PhoZ is only active following its export across the cytoplasmic membrane, a simple and rapid colorimetric screen can be used to identify colonies arising from a mutant that contains an insertion in a gene that encodes a secreted protein. The Tn*FuZ* element has been used to identify secreted proteins whose expression is enhanced during aerobic growth (81). Another example of the versatility of IS*256*-based elements is those that have been modified to include genes encoding a bacterial luciferase (use of luciferase as a reporter gene is discussed below). Insertion into a transcription unit that directs transcription of the luciferase genes allows gene expression *in vivo* to be monitored noninvasively using specialized imaging cameras that can capture photons emitted by the streptococci in living host tissue (82).

Transposome Mutagenesis

In transposome mutagenesis, a stable complex between transposase and transposon DNA is formed *in vitro* and then introduced into a bacterium. The elevated concentration of Mg^{2+} in the bacterial cytoplasm activates transposase that then catalyzes transposition of the element to some random location in the bacterial chromosome. This method has been adapted for efficient mutagenesis of *S. pyogenes* (43). For this, a streptococcal erythromycin-resistance gene was introduced between the inverted repeats of a derivative of Tn*5* (EZ-TN5^{tm}) on the *E. coli* plasmid pMOD-2 (Epicentre Technologies) to construct TmErm (Fig. 7E) (43). Following digestion with a restriction endonuclease, the transposon (TmErm)-containing linear DNA fragment is reacted with a modified hyperactive transposase (EZ::TN, Epicen-

tre Technologies) to generate the transposome. Following transformation of *S. pyogenes* using electroporation, erythromycin-resistant colonies were obtained at frequencies between 10^3 and 10^4 CFU/μg DNA (43). Mutations generated by this method have proven to be very stable even in the absence of erythromycin selection, possibly because the inserted transposon does not encode its own transposase.

Himar1-Based Mariner Transposon Systems

While the generation of a complex transposon library is relatively straightforward, mapping the location of the insertion site in the chromosome of any single mutant has historically been quite laborious. Prior to the development of rapid and relatively low-cost whole-genome sequencing technology, arbitrary PCR (78) was usually utilized to map the site of transposon insertion. An adaptation of this method for rapid screening using pools of mutants was a form of microarray-based analysis known as transposon site hybridization, utilizing a transposon with outward-facing T7 promoters at flanking ends (pOSKAR) (Fig. 7F) (83). Later, this transposon platform was modified to allow for analysis by high-throughput next-generation sequencing through the use of an outward-cutting restriction enzyme known as *Mme*I for the technique known as TnSeq (pKRMIT) (Fig. 7F) (84). Both of these transposon platforms utilize the highly active and promiscuous *mariner Himar1* transposable element, whose target insertion site requires only a thiamine-adenine (TA) dinucleotide sequence, leading to a highly random library of insertions.

Le Breton and McIver's group developed a Mariner-based system for use in *S. pyogenes* by introducing the *mariner* transposon into the ts pWV01 shuttle vector. Because the transposon is highly active, propagation of the shuttle vector in *E. coli* is unstable and requires careful screening of preparations of purified plasmid to ensure that the correct plasmid structure has been maintained prior to electroporation into *S. pyogenes* (12). Furthermore, transposition likely occurs quite rapidly relative to the rate of plasmid loss following growth at restrictive plasmid temperatures. Therefore, additional screening is required for the subsequent *S. pyogenes* mutant pools prior to phenotypic testing. Despite these technical limitations, multiple labs have used this transposon system for transposon site hybridization analysis utilizing the plasmid pOSKAR to screen for genes required for growth in human blood (25), to identify regulators of the Rgg quorum sensing system (85, 86), and for TnSeq analysis utilizing the plasmid pKRMIT for analysis of core essential genes required for growth *in vitro* (42) and for murine soft-tissue infection (87).

Analysis of Transposon Mutants

Regardless of the type of transposable element utilized, there are a number of additional tests that should be performed to ensure that the mutant phenotype is the direct result of insertion of the transposon. This involves determining the sequence of the locus into which the transposon has inserted via arbitrary PCR, inverse PCR after enzyme digestion and ligation, or whole-genome sequencing. With this data in hand, the next step is to construct a mutant in the locus in the wild-type parental strain using one of the strategies for directed mutagenesis outlined above. If the original insertion is responsible for the mutant phenotype, then it is expected that all isolates obtained via directed mutagenesis will also be mutants. It should also be kept in mind when interpreting results that transposon-generated mutations are also polar, so the gene(s) responsible for the observed phenotype may lie downstream of the interrupted gene.

ANALYSIS OF GENE EXPRESSION

Gene regulation phenomena play a key role in pathogenesis. The interaction between the host and the microbe is dynamic and often is a progression through a number of discrete steps. Each of these steps is characterized by the expression of specific sets of bacterial genes required for survival and multiplication and the host's response to the action of the products of these microbial genes. As a consequence, the microorganism is continually challenged to adapt to new and changing environmental conditions. Pathogens have taken advantage of the dynamic nature of this interaction and have evolved to recognize changes in specific environmental conditions as markers that define a particular host compartment or stage of infection. The pathogen uses this information to modulate expression of virulence genes required for survival in this host compartment. Thus, an understanding of the *in vitro* conditions that regulate expression of a specific virulence gene provides insight into how the gene contributes to virulence *in vivo* (for review see 88). Regulation of virulence genes often involves control at the level of transcription. For *S. pyogenes*, transcriptional regulation has often been analyzed through the use of reporter genes, particularly in studies designed to determine the specific conditions that activate or repress expression of a gene of interest.

The basic strategy for using reporter genes to analyze gene expression involves fusing the promoter for the virulence gene of interest to a reporter gene that encodes a gene whose product can easily be quantitated. This feature is of particular utility when the assays for quan-

titation of the product of the streptococcal gene under analysis are time-consuming, expensive, and/or cumbersome. The reporter gene lacks its own promoter, so its product becomes an accurate relative indicator of the steady state level of transcription initiation from the target promoter. It should be kept in mind that because the message is a chimera between the initiation signals of the target gene and the reporter gene, it is not subject to the same posttranscriptional and posttranslational controls. Furthermore, the half-life of the reporter gene's translation product does not reflect that of the native polypeptide. Thus, reporter genes are most appropriately used to quantitate the strength of initiation of transcription of the target promoter.

The reporter genes that have been successfully used in *S. pyogenes* include chloramphenicol acetyltransferase (Cat), β-galactosidase (Lac), β-glucuronidase (Gus), alkaline phosphatase (Pho), and bacterial luciferase (LuxAB). The products of all these genes are stable, highly active enzymes that are highly specific for their substrates. Their utility is enhanced by the availability of synthetic substrates with isotopic, colorimeteric, fluorogenic, or light-emitting properties that allow very sensitive detection. Also, the reporter genes encoding these enzymes are derived from Gram-positive bacteria or have been extensively modified to optimize their expression in Gram-positive hosts and differ extensively from their counterparts used for analysis in *E. coli*.

The most successful applications of Cat have employed *cat86*, originally isolated from *Bacillus pumilus*, which has been modified by the deletion of an attenuator sequence and the substitution of an adenine-thiamine-guanine (ATG) start codon for the native gene's thiamine-thiamine-guanine (TTG) start codon. The advantages of Cat are that *cat86* is very stably maintained in *S. pyogenes* and that the assays for its enzymatic acetylation of chloramphenicol are very sensitive (89, 90). It has several disadvantages: the most sensitive assays involve radioactivity, are expensive, are somewhat labor-intensive relative to other reporter genes, and require the preparation of cell-free extracts. Both Lac and Gus utilize *E. coli* genes that have been modified by the introduction of ribosome binding sites recognized by Gram-positive bacteria (91, 92). The advantages to using these as reporters are that assays for their enzymatic activities are easy to perform and rapid and many samples can be analyzed simultaneously, that a wide variety of substrates is available, and that it is possible to analyze both permeabilized cells and cell extracts. Disadvantages are that they are not as sensitive or as stably maintained in *S. pyogenes* as *cat86*. Nevertheless, *gusA* has proven to be very useful in numerous applications, including identification of regulatory sites in the *hasA* promoter (93) and to find a regulator (*rocA*) that modulates transcription of the CovRS two-component global regulator of virulence and stress resistance (94).

The LuxAB bacterial luciferase, originally from *Vibrio fischeri*, has found utility in dissecting the mechanism of Rgg-associated signaling peptides in *S. pyogenes* (35). An advantage of a luciferase reporter is that the assay is extremely sensitive and quantitative, because exposure of bacterial cultures to volatile decyl aldehyde (decanal) for less than one minute can result in a measurable light output. A disadvantage of this reporter is that luciferase activity is growth phase sensitive due to the reaction's requirement for reduced flavin mononucleotide as a cofactor (95) and has been shown to fail in some bacteria during the stationary phase (82) or under conditions of low pH or oxidative stress (96).

A chimeric reporter protein based on fusion of two naturally secreted proteins has also been developed. This reporter contains as its enzymatic partner, the alkaline phosphatase of *E. faecalis* (PhoZ) (97). This enzyme is secreted from *E. faecalis* and is a lipoprotein that becomes tethered to the outer leaflet of the cell membrane. Fusing the enzymatic C-terminal domain of PhoZ, not including its lipoprotein secretion signal, to the nonlipoprotein N-terminal secretion domain of protein F of *S. pyogenes* results in a very stable, highly enzymatically active chimeric protein (PhoZF) (Fig. 1) which is freely secreted from the streptococcal cell into the surrounding medium (20). This feature makes quantitative analysis of the secreted chimeric protein in culture supernatant very simple and requires neither permeabilization nor preparation of cytoplasmic extracts. Due to the widespread popularity of alkaline phosphatase enzymes, an excellent selection of sensitive substrates with a variety of useful characteristics, including those with fluorescent and light-emitting properties, are inexpensive and are readily available.

Green fluorescent protein (GFP) has become a widely used reporter of gene expression because of its ease of detection and its utility for application in real-time assays for analyses *in vitro* and *in vivo* during infection. Numerous versions of GFP are available that have been modified to alter codon usage, stability, and spectral properties (98). The GFPuv variant has been used to examine regulation of the promoter for the hyaluronic acid biosynthesis gene cluster in streptococci recovered directly from infected tissues (99). This latter study also reported an important consideration for interpreting data derived from GFP-based reporters in that fluorescence is often not uniform among streptococcal cells in

any population, including many cells that express no detectable fluorescence (99). It is likely that the microaerophilic or anaerobic conditions that are typically used for culturing *S. pyogenes* are not ideal for GFP maturation, since oxygen is required for GFP chromophore formation. However, a more stable variant known as superfolder GFP has shown utility in several applications in *S. pyogenes*, including monitoring subcellular localization of proteins (100).

Numerous strategies have been used to introduce reporter fusion constructs into *S. pyogenes* for analysis. Plasmid vectors have the advantage that they are easy to manipulate and can be used to generate and analyze a large number of permutations and variations of promoter structure to probe the *cis* and *trans*-acting control regions. A plasmid-based reporter is present in multiple copies, and this amplification increases the sensitivity of detecting the activity of the promoter under analysis. This latter feature can also make plasmid-based systems sensitive to multiple-copy-number-derived artifacts, such as titration of *trans*-acting regulatory components, when those components are only present in limiting quantities. Furthermore, any sequence-directed local effects of chromosomal structure on expression of a given promoter will not likely be replicated in a plasmid environment. Integration of reporter constructs into the chromosome can alleviate many of these latter types of potential artifacts but with a trade-off in ease of use and sensitivity. Integration is usually performed by introducing the reporter construct on a nonreplicating *E. coli*-based plasmid to target integration into the target gene's native locus by homologous recombination promoted by the cloned promoter segment. Alternatively, use of the integrating p7INT/p7INT* plasmid can reproduce a more native chromosomal environment through integration at the *attB* site (Fig. 2).

Next-generation sequencing approaches such as RNA seq for global analysis of the transcriptome coupled with quantitative real-time RT-PCR methods have recently become the standard for analysis of gene expression in *S. pyogenes*. Although it requires preexisting knowledge of the genome sequence, this disadvantage is more than offset by the facts that RNAseq does not require the laborious construction of reporter gene fusions, that it is fast and easy to perform, that it can directly analyze the entire transcriptome, including novel RNA species, and that it is not limited to the genes for which probes were designed. RNA analysis has been greatly facilitated by the development of technology for the simple and rapid purification of high-quality RNA from relatively small samples (101). In practical terms, quantitative real-time RT-PCR is used for validation expression of individual genes of interest from data developed from RNAseq analyses (102). In quantitative real-time RT-PCR, total RNA is purified and then converted to first-strand complementary DNA (cDNA) using reverse transcriptase. These cDNAs then serve as a template for real-time PCR using primers specific for the gene in question. Real-time thermocyclers are modified with a detection system that can continuously monitor the fluorescence of a probe that is sensitive to the quantity of DNA that is being synthesized by PCR, and the amplification cycle at which a detectable amplification product is first detected is a quantitative measure of the concentration of that message in the input cDNA pool. Data are typically reported as a ratio to a reference gene such as *recA* or *gyrA* whose expression is relatively constant under analysis conditions.

Techniques for Ectopic Expression

Structure-function studies of the role of putative virulence factors require techniques for ectopic expression, both to control the timing and relative levels of expression of a given gene and to study the structure-function relationships of specific domains of a given virulence-associated protein. The former case requires the availability of an ectopic promoter whose activity can be tightly, easily, and quantitatively controlled by some external factor. The latter case requires a method for the expression of various domains of the polypeptide under analysis into the context of an unrelated polypeptide. This allows an independent analysis of the functionality of that domain separate from other regions of the original polypeptide. Methods to accomplish both of these expression strategies have been developed for analysis of virulence in *S. pyogenes*.

An initial approach for using a regulated heterologous promoter to direct expression of the gene under analysis was based on the *nisA* promoter of *Lactococcus lactis*. In lactococci, the NisR/K two-component system controls the expression PnisA in direct response to nisin concentration. The method for using this regulated promoter involves the introduction of *nisK* and *nisR* into an *S. pyogenes* strain, while the gene of interest is introduced under control of the *nisA* promoter, which makes expression of the gene of interest sensitive to the concentration of exogenous nisin that is added to the culture. Tightly regulated expression of a *gusA* reporter gene and induction of up to 60-fold over background using a broad range of nisin concentrations has been reported for *S. pyogenes* (74). The NisR/K system has some drawbacks, including the stability of nisin (103) and a relatively high basal level of expression in the uninduced state, which makes conditional expression of putative essential

genes difficult. A more tightly controlled inducible promoter is based on the tetracycline-inducible promoter P_{tet}. In this system, the *tet* repressor (TetR), binds to a 19-nucleotide palindromic sequence, referred to as the *tet* operator (*tetO)*, between the -35 and -10 site of P_{tet}. In the presence of inducer, commonly tetracycline or nonantimicrobial analogs, including anhydrotetracycline (aTc), TetR binds to the inducer, reducing its binding affinity for *tetO* and thereby allowing for expression of the downstream open reading frame. Originally from *E. coli*, this system has been adapted for use in Gram-positive bacteria by the use of optimized promoter sequences and ribosome-binding sites and through addition of multiple *tetO* sites for enhanced dynamic range (104). A reverse TetR has been developed that has an enhanced, rather than reduced, affinity for *tetO* in the presence of inducer, and thereby the presence of inducer represses gene expression (105). P_{tet} has been used in *S. pyogenes* to control the expression of reporter genes (100) and to experimentally verify the essentiality of ribonucleases J1 and J2 (106). A third, and perhaps more widely used, conditional expression system is based on a theophylline-inducible riboswitch (32). In this system, a ribosome-binding site (RBS) for expression of the target gene has been engineered to become occluded and unavailable due to mRNA secondary structure. However, the binding of the theophylline ligand to the mRNA induces the formation of an alternative secondary structure, such that the RBS is free to bind ribosomes and translation is initiated. The native RBS located upstream of a gene of interest is thus replaced by the riboswitch for conditional expression studies. Two advantages of the theophylline-inducible riboswitch are less leaky expression than P_{tet} and the fact that riboswitches do not require accessory proteins. The theophylline riboswitch has been used to study the essentiality of the histone-like protein gene *hlp* (107), *gacA* (108), and several genes predicted to be essential from a saturating TNseq transposon screen (42).

Structure-function dissection of virulence protein function can benefit from ectopic expression methods that allow the analysis of defined subdomains of the polypeptide. This has proven particularly valuable for identification of the ligand-binding domains of adhesins that direct attachment of bacteria to various host cell receptors. In this approach, the subdomain under examination is expressed and displayed on the bacterial cell surface in the context of a heterologous protein. A technique developed for this strategy, which uses the framework of the M6.1 protein, was based on the work of Pozzi and coworkers (109), who demonstrated that the central domain of the M protein could be replaced with a heterologous sequence without affecting expression and surface presentation of the resulting chimera. Essentially, the heterologous domain is placed between the N-terminal secretion and C-terminal attachment domains of M protein to display the heterologous domain on the streptococcal cell surface. A successful application of this strategy involved analysis of the two distinct fibronectin-binding domains of protein F (110). For a more detailed explanation of this technique see Hanski et al. (39).

CONCLUDING REMARKS

Rapid progress has been made in recent years in the development of sophisticated techniques for genetic analysis in *S. pyogenes*. Much of this effort has been directed at the development of methods for the mutagenesis of known genes. Considerable progress has also been made in the development of strategies for the identification of novel genes. It is likely that the widespread application of these techniques to the virulence properties of *S. pyogenes* will enrich our understanding of streptococcal pathogenesis with insight at the molecular level and will help to establish and clarify the contributions of specific genes. Additional use and development of methods for the analysis of gene expression and heterologous expression will continue and will allow analyses of virulence factors at much higher levels of resolution than previously possible.

REFERENCES

1. **Falkow S.** 1988. Molecular Koch's postulates applied to microbial pathogenicity. *Rev Infect Dis* **10**(Suppl 2):S274–S276 http://dx.doi.org/10.1093/cid/10.Supplement_2.S274.

2. **Ferretti JJ, McShan WM, Ajdic D, Savic DJ, Savic G, Lyon K, Primeaux C, Sezate S, Suvorov AN, Kenton S, Lai HS, Lin SP, Qian Y, Jia HG, Najar FZ, Ren Q, Zhu H, Song L, White J, Yuan X, Clifton SW, Roe BA, McLaughlin R.** 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A* **98**:4658–4663 http://dx.doi.org/10.1073/pnas.071559398.

3. **NCBI.** *Streptococcus pyogenes*. https://www.ncbi.nlm.nih.gov/genome/genomes/175.

4. **Maruyama F, Watanabe T, Nakagawa I.** 2016. *Streptococcus pyogenes* genomics. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), *Streptococcus pyogenes: Basic Biology to Clinical Manifestations*. University of Oklahoma Health Sciences Center, Oklahoma City, OK.

5. **Bessen DE, McShan WM, Nguyen SV, Shetty A, Agrawal S, Tettelin H.** 2015. Molecular epidemiology and genomics of group A *Streptococcus*. *Infect Genet Evol* **33**:393–418 http://dx.doi.org/10.1016/j.meegid.2014.10.011.

6. **Rago JV, Schlievert PM.** 1998. Mechanisms of pathogenesis of staphylococcal and streptococcal superantigens. *Curr Top Microbiol Immunol* **225**:81–97 http://dx.doi.org/10.1007/978-3-642-80451-9_5.

7. **Caparon MG, Scott JR.** 1987. Identification of a gene that regulates expression of M protein, the major virulence determinant of group A streptococci. *Proc Natl Acad Sci U S A* **84**:8677–8681 http://dx.doi.org/10.1073/pnas.84.23.8677.

8. Moynet DJ, Colon-Whitt AE, Calandra GB, Cole RM. 1985. Structure of eight streptococcal bacteriophages. *Virology* **142:**263–269 http://dx.doi.org/10.1016/0042-6822(85)90334-4.

9. Caparon MG, Scott JR. 1991. Genetic manipulation of pathogenic streptococci. *Methods Enzymol* **204:**556–586 http://dx.doi.org/10.1016/0076-6879(91)04028-M.

10. Malke H. 1969. Transduction of *Streptococcus pyogenes* K 56 by temperature-sensitive mutants of the transducing phage A 25. *Z Naturforsch B* **24:**1556–1561 http://dx.doi.org/10.1515/znb-1969-1214.

11. Dunny GM, Lee LN, LeBlanc DJ. 1991. Improved electroporation and cloning vector system for Gram-positive bacteria. *Appl Environ Microbiol* **57:**1194–1201.

12. Le Breton Y, McIver KS. 2013. Genetic manipulation of *Streptococcus pyogenes* (the group A streptococcus, GAS). *Curr Protoc Microbiol* **30:**9D.3.1–9D.3.29.

13. Husmann LK, Yung DL, Hollingshead SK, Scott JR. 1997. Role of putative virulence factors of *Streptococcus pyogenes* in mouse models of long-term throat colonization and pneumonia. *Infect Immun* **65:**1422–1430.

14. Simon D, Ferretti JJ. 1991. Electrotransformation of *Streptococcus pyogenes* with plasmid and linear DNA. *FEMS Microbiol Lett* **66:**219–224 http://dx.doi.org/10.1111/j.1574-6968.1991.tb04868.x.

15. Mashburn-Warren L, Morrison DA, Federle MJ. 2012. The cryptic competence pathway in *Streptococcus pyogenes* is controlled by a peptide pheromone. *J Bacteriol* **194:**4589–4600 http://dx.doi.org/10.1128/JB.00830-12.

16. Ji Y, McLandsborough L, Kondagunta A, Cleary PP. 1996. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. *Infect Immun* **64:**503–510.

17. Hidalgo-Grass C, Ravins M, Dan-Goor M, Jaffe J, Moses AE, Hanski E. 2002. A locus of group A *Streptococcus* involved in invasive disease and DNA transfer. *Mol Microbiol* **46:**87–99 http://dx.doi.org/10.1046/j.1365-2958.2002.03127.x.

18. Leenhouts KJ, Tolner B, Bron S, Kok J, Venema G, Seegers JF. 1991. Nucleotide sequence and characterization of the broad-host-range lactococcal plasmid pWVO1. *Plasmid* **26:**55–66 http://dx.doi.org/10.1016/0147-619X(91)90036-V.

19. Biswas I, Jha JK, Fromm N. 2008. Shuttle expression plasmids for genetic studies in *Streptococcus mutans*. *Microbiology* **154:**2275–2282 http://dx.doi.org/10.1099/mic.0.2008/019265-0.

20. Granok AB, Parsonage D, Ross RP, Caparon MG. 2000. The RofA binding site in *Streptococcus pyogenes* is utilized in multiple transcriptional pathways. *J Bacteriol* **182:**1529–1540 http://dx.doi.org/10.1128/JB.182.6.1529-1540.2000.

21. Bryksin AV, Matsumura I. 2010. Rational design of a plasmid origin that replicates efficiently in both Gram-positive and Gram-negative bacteria. *PLoS One* **5:**e13244 http://dx.doi.org/10.1371/journal.pone.0013244.

22. Maguin E, Duwat P, Hege T, Ehrlich D, Gruss A. 1992. New thermosensitive plasmid for Gram-positive bacteria. *J Bacteriol* **174:**5633–5638 http://dx.doi.org/10.1128/jb.174.17.5633-5638.1992.

23. Perez-Casal J, Price JA, Maguin E, Scott JR. 1993. An M protein with a single C repeat prevents phagocytosis of *Streptococcus pyogenes*: use of a temperature-sensitive shuttle vector to deliver homologous sequences to the chromosome of *S. pyogenes*. *Mol Microbiol* **8:**809–819 http://dx.doi.org/10.1111/j.1365-2958.1993.tb01628.x.

24. Nielsen HV, Guiton PS, Kline KA, Port GC, Pinkner JS, Neiers F, Normark S, Henriques-Normark B, Caparon MG, Hultgren SJ. 2012. The metal ion-dependent adhesion site motif of the *Enterococcus faecalis* EbpA pilin mediates pilus function in catheter-associated urinary tract infection. *MBio* **3:**e00177-12 http://dx.doi.org/10.1128/mBio.00177-12.

25. Le Breton Y, Mistry P, Valdes KM, Quigley J, Kumar N, Tettelin H, McIver KS. 2013. Genome-wide identification of genes required for fitness of group A *Streptococcus* in human blood. *Infect Immun* **81:**862–875 http://dx.doi.org/10.1128/IAI.00837-12.

26. Scott JR, Guenthner PC, Malone LM, Fischetti VA. 1986. Conversion of an M^- group A streptococcus to M^+ by transfer of a plasmid containing an M6 gene. *J Exp Med* **164:**1641–1651 http://dx.doi.org/10.1084/jem.164.5.1641.

27. Hanski E, Horwitz PA, Caparon MG. 1992. Expression of protein F, the fibronectin-binding protein of *Streptococcus pyogenes* JRS4, in heterologous streptococcal and enterococcal strains promotes their adherence to respiratory epithelial cells. *Infect Immun* **60:**5119–5125.

28. DeAngelis PL, Papaconstantinou J, Weigel PH. 1993. Isolation of a *Streptococcus pyogenes* gene locus that directs hyaluronan biosynthesis in acapsular mutants and in heterologous bacteria. *J Biol Chem* **268:**14568–14571.

29. Clewell DB, Yagi Y, Dunny GM, Schultz SK. 1974. Characterization of three plasmid deoxyribonucleic acid molecules in a strain of *Streptococcus faecalis*: identification of a plasmid determining erythromycin resistance. *J Bacteriol* **117:**283–289.

30. Trieu-Cuot P, Carlier C, Poyart-Salmeron C, Courvalin P. 1990. A pair of mobilizable shuttle vectors conferring resistance to spectinomycin for molecular cloning in *Escherichia coli* and in Gram-positive bacteria. *Nucleic Acids Res* **18:**4296 http://dx.doi.org/10.1093/nar/18.14.4296.

31. Fogg GC, Caparon MG. 1997. Constitutive expression of fibronectin binding in *Streptococcus pyogenes* as a result of anaerobic activation of rofA. *J Bacteriol* **179:**6172–6180 http://dx.doi.org/10.1128/jb.179.19.6172-6180.1997.

32. Topp S, Reynoso CM, Seeliger JC, Goldlust IS, Desai SK, Murat D, Shen A, Puri AW, Komeili A, Bertozzi CR, Scott JR, Gallivan JP. 2010. Synthetic riboswitches that induce gene expression in diverse bacterial species. *Appl Environ Microbiol* **76:**7881–7884 http://dx.doi.org/10.1128/AEM.01537-10.

33. McShan WM, McLaughlin RE, Nordstrand A, Ferretti JJ. 1998. Vectors containing streptococcal bacteriophage integrases for site-specific gene insertion. *Methods Cell Sci* **20:**51–57 http://dx.doi.org/10.1023/A:1009773309163.

34. Shulman M, Gottesman M. 1973. Attachment site mutants of bacteriophage lambda. *J Mol Biol* **81:**461–482 http://dx.doi.org/10.1016/0022-2836(73)90517-2.

35. Chang JC, LaSarre B, Jimenez JC, Aggarwal C, Federle MJ. 2011. Two group A streptococcal peptide pheromones act through opposing Rgg regulators to control biofilm development. *PLoS Pathog* **7:**e1002190 http://dx.doi.org/10.1371/journal.ppat.1002190.

36. Perez-Casal J, Caparon MG, Scott JR. 1991. Mry, a trans-acting positive regulator of the M protein gene of *Streptococcus pyogenes* with similarity to the receptor proteins of two-component regulatory systems. *J Bacteriol* **173:**2617–2624 http://dx.doi.org/10.1128/jb.173.8.2617-2624.1991.

37. Tesorero RA, Yu N, Wright JO, Svencionis JP, Cheng Q, Kim JH, Cho KH. 2013. Novel regulatory small RNAs in *Streptococcus pyogenes*. *PLoS One* **8:**e64021 http://dx.doi.org/10.1371/journal.pone.0064021.

38. Fogg GC, Gibson CM, Caparon MG. 1994. The identification of rofA, a positive-acting regulatory component of prtF expression: use of an m gamma delta-based shuttle mutagenesis strategy in *Streptococcus pyogenes*. *Mol Microbiol* **11:**671–684 http://dx.doi.org/10.1111/j.1365-2958.1994.tb00345.x.

39. Hanski E, Fogg G, Tovi A, Okada N, Burstein I, Caparon M. 1995. Molecular analysis of *Streptococcus pyogenes* adhesion. *Methods Enzymol* **253:**269–305 http://dx.doi.org/10.1016/S0076-6879(95)53025-8.

40. Okada N, Geist RT, Caparon MG. 1993. Positive transcriptional control of *mry* regulates virulence in the group A streptococcus. *Mol Microbiol* **7:**893–903 http://dx.doi.org/10.1111/j.1365-2958.1993.tb01180.x.

41. Cho KH, Caparon MG. 2005. Patterns of virulence gene expression differ between biofilm and tissue communities of *Streptococcus pyogenes*. *Mol Microbiol* **57:**1545–1556 http://dx.doi.org/10.1111/j.1365-2958.2005.04786.x.

42. Le Breton Y, Belew AT, Valdes KM, Islam E, Curry P, Tettelin H, Shirtliff ME, El-Sayed NM, McIver KS. 2015. Essential genes in the core genome of the human pathogen *Streptococcus pyogenes*. *Sci Rep* **5:**9838 http://dx.doi.org/10.1038/srep09838.

43. Cho KH, Caparon MG. 2008. tRNA modification by GidA/MnmE is necessary for *Streptococcus pyogenes* virulence: a new strategy to make live attenuated strains. *Infect Immun* **76:**3176–3186 http://dx.doi.org/10.1128/IAI.01721-07.

44. Moldover B, Piggot PJ, Yudkin MD. 1991. Identification of the promoter and the transcriptional start site of the spoVA operon of *Bacillus subtilis* and *Bacillus licheniformis*. *J Gen Microbiol* **137:**527–531 http://dx.doi.org/10.1099/00221287-137-3-527.

45. Port GC, Cusumano ZT, Tumminello PR, Caparon MG. 2017. SpxA1 and SpxA2 act coordinately to fine-tune stress responses and virulence in *Streptococcus pyogenes*. *MBio* **8:**e00288-17 http://dx.doi.org/10.1128/mBio.00288-17.

46. Ashbaugh CD, Warren HB, Carey VJ, Wessels MR. 1998. Molecular analysis of the role of the group A streptococcal cysteine protease, hyaluronic acid capsule, and M protein in a murine model of human invasive soft-tissue infection. *J Clin Invest* **102:**550–560 http://dx.doi.org/10.1172/JCI3065.

47. Ruiz N, Wang B, Pentland A, Caparon M. 1998. Streptolysin O and adherence synergistically modulate proinflammatory responses of keratinocytes to group A streptococci. *Mol Microbiol* **27:**337–346 http://dx.doi.org/10.1046/j.1365-2958.1998.00681.x.

48. Horton RM, Cai Z, Ho SM, Pease LR. 2013. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* **54:**129–133 http://dx.doi.org/10.2144/000114017.

49. Bryksin AV, Matsumura I. 2010. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques* **48:**463–465 http://dx.doi.org/10.2144/000113418.

50. Schweizer H. 2008. Bacterial genetics: past achievements, present state of the field, and future challenges. *Biotechniques* **44:**633–634, 636–641 http://dx.doi.org/10.2144/000112807.

51. Kristich CJ, Chandler JR, Dunny GM. 2007. Development of a host-genotype-independent counterselectable marker and a high-frequency conjugative delivery system and their use in genetic analysis of *Enterococcus faecalis*. *Plasmid* **57:**131–144 http://dx.doi.org/10.1016/j.plasmid.2006.08.003.

52. Miyazaki K. 2015. Molecular engineering of a PheS counterselection marker for improved operating efficiency in *Escherichia coli*. *Biotechniques* **58:**86–88 http://dx.doi.org/10.2144/000114257.

53. Zhang S, Zou Z, Kreth J, Merritt J. 2017. Recombineering in *Streptococcus mutans* using direct repeat-mediated cloning-independent markerless mutagenesis (DR-CIMM). *Front Cell Infect Microbiol* **7:**202 http://dx.doi.org/10.3389/fcimb.2017.00202.

54. McIver KS, Myles RL. 2002. Two DNA-binding domains of Mga are required for virulence gene activation in the group A streptococcus. *Mol Microbiol* **43:**1591–1601 http://dx.doi.org/10.1046/j.1365-2958.2002.02849.x.

55. Hille F, Richter H, Wong SP, Bratovič M, Ressel S, Charpentier E. 2018. The biology of CRISPR-Cas: backward and forward. *Cell* **172:**1239–1259 http://dx.doi.org/10.1016/j.cell.2017.11.032.

56. Marraffini LA. 2016. The CRISPR-Cas system of *Streptococcus pyogenes*: function and applications. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), Streptococcus pyogenes: *Basic Biology to Clinical Manifestations*. University of Oklahoma Health Sciences Center, Oklahoma City, OK.

57. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337:**816–821 http://dx.doi.org/10.1126/science.1225829.

58. Xiong X, Chen M, Lim WA, Zhao D, Qi LS. 2016. CRISPR/Cas9 for human genome engineering and disease research. *Annu Rev Genomics Hum Genet* **17:**131–154 http://dx.doi.org/10.1146/annurev-genom-083115-022258.

59. Choi KR, Lee SY. 2016. CRISPR technologies for bacterial systems: current achievements and future directions. *Biotechnol Adv* **34:**1180–1209 http://dx.doi.org/10.1016/j.biotechadv.2016.08.002.

60. Sonoda E, Hochegger H, Saberi A, Taniguchi Y, Takeda S. 2006. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair (Amst)* **5:**1021–1029 http://dx.doi.org/10.1016/j.dnarep.2006.05.022.

61. Bowater R, Doherty AJ. 2006. Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genet* **2:**e8 http://dx.doi.org/10.1371/journal.pgen.0020008.

62. Chen W, Zhang Y, Yeo WS, Bae T, Ji Q. 2017. Rapid and efficient genome editing in *Staphylococcus aureus* by using an engineered CRISPR/Cas9 System. *J Am Chem Soc* **139:**3790–3795 http://dx.doi.org/10.1021/jacs.6b13317.

63. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* **31:**233–239 http://dx.doi.org/10.1038/nbt.2508.

64. Cho KH, Wright J, Svencionis J, Kim JH. 2013. The prince and the pauper: which one is real? The problem of secondary mutation during mutagenesis in *Streptococcus pyogenes*. *Virulence* **4:**664–665 http://dx.doi.org/10.4161/viru.26767.

65. Ashbaugh CD, Wessels MR. 2001. Absence of a cysteine protease effect on bacterial virulence in two murine models of human invasive group A streptococcal infection. *Infect Immun* **69:**6683–6688 http://dx.doi.org/10.1128/IAI.69.11.6683-6686.2001.

66. Port GC, Vega LA, Nylander AB, Caparon MG. 2014. *Streptococcus pyogenes* polymyxin B-resistant mutants display enhanced ExPortal integrity. *J Bacteriol* **196:**2563–2577 http://dx.doi.org/10.1128/JB.01596-14.

67. Scott JR, Bringel F, Marra D, Van Alstine G, Rudy CK. 1994. Conjugative transposition of Tn*916*: preferred targets and evidence for conjugative transfer of a single strand and for a double-stranded circular intermediate. *Mol Microbiol* **11:**1099–1108 http://dx.doi.org/10.1111/j.1365-2958.1994.tb00386.x.

68. Betschel SD, Borgia SM, Barg NL, Low DE, De Azavedo JCS. 1998. Reduced virulence of group A streptococcal Tn*916* mutants that do not produce streptolysin S. *Infect Immun* **66:**1671–1679.

69. Liu S, Sela S, Cohen G, Jadoun J, Cheung A, Ofek I. 1997. Insertional inactivation of streptolysin S expression is associated with altered riboflavin metabolism in *Streptococcus pyogenes*. *Microb Pathog* **22:**227–234 http://dx.doi.org/10.1006/mpat.1996.0107.

70. Nida K, Cleary PP. 1983. Insertional inactivation of streptolysin S expression in *Streptococcus pyogenes*. *J Bacteriol* **155:**1156–1161.

71. Dougherty BA, van de Rijn I. 1992. Molecular characterization of a locus required for hyaluronic acid capsule production in group A streptococci. *J Exp Med* **175:**1291–1299 http://dx.doi.org/10.1084/jem.175.5.1291.

72. Wessels MR, Moses AE, Goldberg JB, DiCesare TJ. 1991. Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. *Proc Natl Acad Sci U S A* **88:**8317–8321 http://dx.doi.org/10.1073/pnas.88.19.8317.

73. Tomich PK, An FY, Clewell DB. 1980. Properties of erythromycin-inducible transposon Tn*917* in *Streptococcus faecalis*. *J Bacteriol* **141:**1366–1374.

74. Eichenbaum Z, Federle MJ, Marra D, de Vos WM, Kuipers OP, Kleerebezem M, Scott JR. 1998. Use of the lactococcal *nisA* promoter to regulate gene expression in Gram-positive bacteria: comparison of induction level and promoter strength. *Appl Environ Microbiol* **64:**2763–2769.

75. Eichenbaum Z, Scott JR. 1997. Use of Tn*917* to generate insertion mutations in the group A streptococcus. *Gene* **186:**213–217 http://dx.doi.org/10.1016/S0378-1119(96)00705-6.

76. Camilli A, Portnoy A, Youngman P. 1990. Insertional mutagenesis of *Listeria monocytogenes* with a novel Tn*917* derivative that allows direct cloning of DNA flanking transposon insertions. *J Bacteriol* **172:**3738–3744 http://dx.doi.org/10.1128/jb.172.7.3738-3744.1990.

77. Lyon BR, May JW, Skurray RA. 1984. Tn*4001*: a gentamicin and kanamycin resistance transposon in *Staphylococcus aureus*. *Mol Gen Genet* **193:**554–556 http://dx.doi.org/10.1007/BF00382099.

78. Lyon WR, Gibson CM, Caparon MG. 1998. A role for trigger factor and an *rgg*-like regulator in the transcription, secretion and processing of the cysteine proteinase of *Streptococcus pyogenes*. *EMBO J* **17:**6263–6275 http://dx.doi.org/10.1093/emboj/17.21.6263.

79. Kizy AE, Neely MN. 2009. First *Streptococcus pyogenes* signature-tagged mutagenesis screen identifies novel virulence determinants. *Infect Immun* **77:**1854–1865 http://dx.doi.org/10.1128/IAI.01306-08.

80. Miller JD, Neely MN. 2005. Large-scale screen highlights the importance of capsule for virulence in the zoonotic pathogen *Streptococcus iniae*. *Infect Immun* **73:**921–934 http://dx.doi.org/10.1128/IAI.73.2.921-934.2005.

81. Gibson CM, Caparon MG. 2002. Alkaline phosphatase reporter transposon for identification of genes encoding secreted proteins in Gram-positive microorganisms. *Appl Environ Microbiol* **68:**928–932 http://dx.doi.org/10.1128/AEM.68.02.928-932.2002.

82. Francis KP, Yu J, Bellinger-Kawahara C, Joh D, Hawkinson MJ, Xiao G, Purchio TF, Caparon MG, Lipsitch M, Contag PR. 2001. Visualizing pneumococcal infections in the lungs of live mice using bioluminescent *Streptococcus pneumoniae* transformed with a novel Gram-positive *lux* transposon. *Infect Immun* **69:**3350–3358 http://dx.doi.org/10.1128/IAI.69.5.3350-3358.2001.

83. Sassetti CM, Boyd DH, Rubin EJ. 2001. Comprehensive identification of conditionally essential genes in mycobacteria. *Proc Natl Acad Sci U S A* **98:**12712–12717 http://dx.doi.org/10.1073/pnas.231275498.

84. van Opijnen T, Bodi KL, Camilli A. 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods* **6:**767–772 http://dx.doi.org/10.1038/nmeth.1377.

85. Chang JC, Federle MJ. 2016. PptAB exports Rgg quorum-sensing peptides in *Streptococcus*. *PLoS One* **11:**e0168461 http://dx.doi.org/10.1371/journal.pone.0168461.

86. Wilkening RV, Chang JC, Federle MJ. 2016. PepO, a CovRS-controlled endopeptidase, disrupts *Streptococcus pyogenes* quorum sensing. *Mol Microbiol* **99:**71–87 http://dx.doi.org/10.1111/mmi.13216.

87. Le Breton Y, Belew AT, Freiberg JA, Sundar GS, Islam E, Lieberman J, Shirtliff ME, Tettelin H, El-Sayed NM, McIver KS. 2017. Genome-wide discovery of novel M1T1 group A streptococcal determinants important for fitness and virulence during soft-tissue infection. *PLoS Pathog* **13:**e1006584 http://dx.doi.org/10.1371/journal.ppat.1006584.

88. Mekalanos JJ. 1992. Environmental signals controlling expression of virulence determinants in bacteria. *J Bacteriol* **174:**1–7 http://dx.doi.org/10.1128/jb.174.1.1-7.1992.

89. Caparon MG, Geist RT, Perez-Casal J, Scott JR. 1992. Environmental regulation of virulence in group A streptococci: transcription of the gene encoding M protein is stimulated by carbon dioxide. *J Bacteriol* **174:**5693–5701 http://dx.doi.org/10.1128/jb.174.17.5693-5701.1992.

90. McIver KS, Heath AS, Scott JR. 1995. Regulation of virulence by environmental signals in group A streptococci: influence of osmolarity, temperature, gas exchange, and iron limitation on *emm* transcription. *Infect Immun* **63:**4540–4542.

91. Ferrari E, Henner DJ, Perego M, Hoch JA. 1988. Transcription of *Bacillus subtilis* subtilisin and expression of subtilisin in sporulation mutants. *J Bacteriol* **170:**289–295 http://dx.doi.org/10.1128/jb.170.1.289-295.1988.

92. Karow ML, Piggot PJ. 1995. Construction of *gusA* transcriptional fusion vectors for *Bacillus subtilis* and their utilization for studies of spore formation. *Gene* **163:**69–74 http://dx.doi.org/10.1016/0378-1119(95)00402-R.

93. Federle MJ, Scott JR. 2002. Identification of binding sites for the group A streptococcal global regulator CovR. *Mol Microbiol* **43:**1161–1172 http://dx.doi.org/10.1046/j.1365-2958.2002.02810.x.

94. Biswas I, Scott JR. 2003. Identification of *rocA*, a positive regulator of *covR* expression in the group A streptococcus. *J Bacteriol* **185:**3081–3090 http://dx.doi.org/10.1128/JB.185.10.3081-3090.2003.

95. Meighen EA. 1991. Molecular biology of bacterial bioluminescence. *Microbiol Rev* **55:**123–142.

96. Wiles S, Ferguson K, Stefanidou M, Young DB, Robertson BD. 2005. Alternative luciferase for monitoring bacterial cells under adverse conditions. *Appl Environ Microbiol* **71:**3427–3432 http://dx.doi.org/10.1128/AEM.71.7.3427-3432.2005.

97. Rothschild CB, Ross RP, Claiborne A. 1991. Molecular analysis of the gene encoding alkaline phophatase in *Streptococcus faecalis* 10C1, p 45–48. *In* Dunny GM, Cleary PP, McKay LL (ed), *Genetics and Molecular Biology of Streptococci, Lactococci and Enterococci*. ASM Press, Washington, DC.

98. Crameri A, Whitehorn EA, Tate E, Stemmer WP. 1996. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol* **14:**315–319 http://dx.doi.org/10.1038/nbt0396-315.

99. Gryllos I, Cywes C, Shearer MH, Cary M, Kennedy RC, Wessels MR. 2001. Regulation of capsule gene expression by group A *Streptococcus* during pharyngeal colonization and invasive infection. *Mol Microbiol* **42:**61–74 http://dx.doi.org/10.1046/j.1365-2958.2001.02635.x.

100. Vega LA, Port GC, Caparon MG. 2013. An association between peptidoglycan synthesis and organization of the *Streptococcus pyogenes* ExPortal. *MBio* **4:**e00485-13 http://dx.doi.org/10.1128/mBio.00485-13.

101. Cheung AL, Eberhardt KJ, Fischetti VA. 1994. A method to isolate RNA from Gram-positive bacteria and mycobacteria. *Anal Biochem* **222:**511–514 http://dx.doi.org/10.1006/abio.1994.1528.

102. Freiberg JA, Le Breton Y, Tran BQ, Scott AJ, Harro JM, Ernst RK, Goo YA, Mongodin EF, Goodlett DR, McIver KS, Shirtliff ME. 2016. Global analysis and comparison of the transcriptomes and proteomes of group A streptococcus biofilms. *mSystems* **1:**00149-16 http://dx.doi.org/10.1128/mSystems.00149-16.

103. Tan Z, Luo J, Liu F, Zhang Q, Jia S. 2015. Effects of pH, temperature, storage time, and protective agents on nisin antibacterial stability. *In* Zhang TC, Nakajima M (ed), *Advances in Applied Biotechnology Lecture Notes in Electrical Engineering*, vol 333. Springer, Berlin, Germany.

104. Geissendörfer M, Hillen W. 1990. Regulated expression of heterologous genes in *Bacillus subtilis* using the Tn*10* encoded *tet* regulatory elements. *Appl Microbiol Biotechnol* **33:**657–663 http://dx.doi.org/10.1007/BF00604933.

105. Kamionka A, Bogdanska-Urbaniak J, Scholz O, Hillen W. 2004. Two mutations in the tetracycline repressor change the inducer anhydrotetracycline to a corepressor. *Nucleic Acids Res* **32:**842–847 http://dx.doi.org/10.1093/nar/gkh200.

106. Bugrysheva JV, Scott JR. 2010. The ribonucleases J1 and J2 are essential for growth and have independent roles in mRNA decay in *Streptococcus pyogenes*. *Mol Microbiol* **75:**731–743 http://dx.doi.org/10.1111/j.1365-2958.2009.07012.x.

107. Bugrysheva JV, Froehlich BJ, Freiberg JA, Scott JR. 2011. The histone-like protein Hlp is essential for growth of *Streptococcus pyogenes*: comparison of genetic approaches to study essential genes. *Appl Environ Microbiol* **77:**4422–4428 http://dx.doi.org/10.1128/AEM.00554-11.

108. van der Beek SL, Le Breton Y, Ferenbach AT, Chapman RN, van Aalten DM, Navratilova I, Boons GJ, McIver KS, van Sorge NM, Dorfmueller HC. 2015. GacA is essential for group A *Streptococcus* and defines a new class of monomeric dTDP-4-dehydrorhamnose reductases (RmlD). *Mol Microbiol* **98:**946–962 http://dx.doi.org/10.1111/mmi.13169.

109. Pozzi G, Contorni M, Oggioni MR, Manganelli R, Tommasino M, Cavalieri F, Fischetti VA. 1992. Delivery and expression of a heterologous antigen on the surface of streptococci. *Infect Immun* **60:**1902–1907.

110. Ozeri V, Tovi A, Burstein I, Natanson-Yaron S, Caparon MG, Yamada KM, Akiyama SK, Vlodavsky I, Hanski E. 1996. A two-domain mechanism for group A streptococcal adherence through protein F to the extracellular matrix. *EMBO J* **15:**989–998 http://dx.doi.org/10.1002/j.1460-2075.1996.tb00435.x.

111. Pritzlaff CA, Chang JC, Kuo SP, Tamura GS, Rubens CE, Nizet V. 2001. Genetic basis for the beta-haemolytic/cytolytic activity of group B *Streptococcus*. *Mol Microbiol* **39:**236–247 http://dx.doi.org/10.1046/j.1365-2958.2001.02211.x.

112. Mashburn-Warren L, Morrison DA, Federle MJ. 2010. A novel double-tryptophan peptide pheromone controls competence in *Streptococcus* spp. via an Rgg regulator. *Mol Microbiol* **78:**589–606 http://dx.doi.org/10.1111/j.1365-2958.2010.07361.x.